Mon. Not. R. Astron. Soc. **320**, 504–528 (2001)

The nature of high-redshift galaxies

Rachel S. Somerville,[1,2]★ Joel R. Primack[2] and S. M. Faber[3]

[1]*Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA*
[2]*Department of Physics, University of California, Santa Cruz, CA 95064, USA*
[3]*UCO/Lick Observatory, University of California, Santa Cruz, CA 95064, USA*

Accepted 2000 August 16. Received 2000 June 29; in original form 1998 June 10

ABSTRACT

Using semi-analytic models of galaxy formation set within the cold dark matter (CDM) merging hierarchy, we investigate several scenarios for the nature of the high-redshift ($z \gtrsim 2$) Lyman-break galaxies (LBGs). We consider a 'collisional starburst' model in which bursts of star formation are triggered by galaxy–galaxy mergers, and find that a significant fraction of LBGs are predicted to be starbursts. This model reproduces the observed comoving number density of bright LBGs as a function of redshift and the observed luminosity function at $z \sim 3$ and $z \sim 4$, with a reasonable amount of dust extinction. Model galaxies at $z \sim 3$ have star formation rates, half-light radii, $I - K$ colours and internal velocity dispersions that are in good agreement with the data. Global quantities such as the star formation rate density and cold gas and metal content of the Universe as a function of redshift also agree well. Two 'quiescent' models without starbursts are also investigated. In one, the star formation efficiency in galaxies remains constant with redshift, while in the other, it scales inversely with disc dynamical time, and thus increases rapidly with redshift. The first quiescent model is strongly ruled out, as it does not produce enough high-redshift galaxies once realistic dust extinction is accounted for. The second quiescent model fits marginally, but underproduces cold gas and very bright galaxies at high redshift. A general conclusion is that star formation at high redshift must be more efficient than locally. The collisional starburst model appears to accomplish this naturally without violating other observational constraints.

Key words: galaxies: evolution – galaxies: formation – galaxies: general – galaxies: high-redshift – galaxies: starburst – cosmology: theory.

1 INTRODUCTION

With the dramatic recent advances in observational astronomy, more and more pieces of the puzzle of galaxy formation and evolution are becoming available. Some of the important puzzle pieces include the number densities, colours, sizes, morphologies, internal velocity dispersions and star formation rates of bright star-forming galaxies spanning a redshift range from $z \sim 0$ to $z \sim 6$, and the complementary information on the neutral hydrogen and metal content of the Universe to $z \sim 4$ obtained from quasar absorption systems. However, it still remains to fit these pieces together into a comprehensive and compelling theoretical framework.

Our window on to the high-redshift ($z \gtrsim 2$) Universe has been expanded tremendously by the 'Lyman-break' photometric selection technique developed by Steidel and collaborators (Steidel & Hamilton 1992, 1993; Steidel et al. 1996a). This technique uses specially developed filters to exploit the redshifted Lyman-limit discontinuity in order to identify high-redshift candidates. Similar techniques were exploited by Madau et al. (1996) to identify high-redshift candidates in the Hubble Deep Field (HDF) (Williams et al. 1996). Extensive spectroscopic follow-up work at the Keck telescope has verified the accuracy of the photometric selection technique (Steidel et al. 1996b, 1999; Lowenthal et al. 1997; Adelberger et al. 1998). The morphologies and sizes of these objects have been studied using the HDF sample (Giavalisco, Steidel & Macchetto 1996; Lowenthal et al. 1997), and their clustering properties have been measured using the growing sample of hundreds (now approaching a thousand) of Lyman-break galaxies (LBGs) with spectroscopically confirmed redshifts (Adelberger et al. 1998; Giavalisco et al. 1998; Steidel et al. 1998). Similar techniques have been used to identify galaxies at $\bar{z} \sim 4$ (Steidel et al. 1999), and a few objects have been discovered with confirmed redshifts $z \gtrsim 5$ (Stern & Spinrad 2000).

Soon after the discovery of Lyman-break galaxies, various interpretations of them were proposed. A view put forward by Steidel et al. (1996a), and reiterated by Giavalisco, Steidel & Macchetto (1996), Adelberger et al. (1998) and Steidel et al.

★ E-mail: rachel@ast.cam.ac.uk

(1998), is that LBGs are located in the centres of massive dark matter haloes ($M \gtrsim 10^{12}\,\mathrm{M}_{\odot}$) and have been forming stars at moderate rates over a fairly long time-scale ($\gtrsim$1 Gyr). This scenario supposes that LBGs are the direct progenitors of today's massive, luminous ellipticals and spheroids, and that such objects had almost completely assembled and begun forming stars at moderate and fairly constant rates by $z \sim 3$. Lowenthal et al. (1997) also equated LBGs with today's massive spheroids, but suggested instead that LBGs are still actively assembling, and that many exist in small-mass haloes raised briefly above the detection limit by short, intense bursts of star formation. Trager et al. (1997) further proposed that these small objects would later be tidally destroyed as they fell into the potential well of larger galaxies to produce a Population II stellar halo like that in the Milky Way.

The impressive body of observations at high redshift would seem to offer a tantalizing promise to provide important constraints on theories of galaxy formation. As usual, the difficulty lies in making a connection between the visible but ill-understood portions of galaxies and the dark matter haloes that are hidden from view but well modelled within the cold dark matter (CDM) framework. Building on the above massive/early-assembly picture for LBGs, several workers investigated a model in which each LBG is associated with a dark matter halo and luminosity scales tightly with halo mass (Mo & Fukugita 1996; Adelberger et al. 1998; Jing & Suto 1998; Wechsler et al. 1998). This work has demonstrated that, in such a picture, the observed number density and clustering of LBGs at $z \sim 3$ can be simultaneously accounted for if there is approximately one galaxy per massive halo ($M \gtrsim 10^{12}\,\mathrm{M}_{\odot}$, where this mass depends somewhat on the cosmological model). We shall refer to this class of simple models as 'massive halo' models. It arises quite generically from such models that LBGs must be substantially more clustered than the underlying dark matter – in short, bright galaxies at high redshift are highly *biased*. This prediction has been emphasized as a major success of the massive halo picture (e.g. Adelberger et al. 1998).

However, the sharp lower mass cut-off and tight link between halo mass and luminosity in massive halo models are doubtless too simplistic. A step forward in realism is provided by semi-analytic models, in which the infall rate of cooling gas determines the star formation rate, and hence the luminosity. Both are modelled as a function of halo mass, which grows via hierarchical clustering, accompanied by a schematic law for gas cooling as a function of halo gas density. Baugh et al. (1998, hereafter BCFL) showed that semi-analytic models are broadly consistent with the simple massive halo picture, and that they also match the observed abundance and clustering of LBGs at $z \sim 3$ (see also Governato et al. 1998).

While these results are encouraging, there is an important caveat. Agreement with the observed number densities of LBGs is achieved in the above models only if internal dust extinction in LBGs is ignored. Much evidence has now accumulated to the effect that dust in LBGs is in fact highly significant (see Section 2.4). The shortfall in LBG numbers that results when realistic dust extinction is included (which we will quantify) motivates us to explore other approaches – in particular, other modes of star formation. We will show how simply changing this ingredient can dramatically change our predictions about the high-redshift Universe, and discuss which – if any – of the recipes can be ruled out by comparison with the observational data.

We will define 'quiescent' star formation to be the standard mode of star formation that occurs within galactic discs whenever cold gas is present. It is the dominant mode of star formation in the present-day Universe, and there is observational evidence that its efficiency in nearby galaxies is related to the internal properties of galactic discs, apparently either to the surface density or to the dynamical time (Kennicutt 1998). Since both of these properties are slowly varying, the star formation rate under quiescent star formation changes only slowly with time.

We consider also a second mode of star formation in which stars are created with dramatically increased efficiency over relatively short time-scales, termed 'starbursts.' Various physical mechanisms might produce such bursts, but here we will assume that they are triggered by galaxy–galaxy mergers. We term these 'collisional' starbursts, to distinguish them from bursts triggered by other means. Substantial observational and theoretical evidence exists in support of the collisional starburst phenomenon. For example, a strong correlation is observed between starburst activity and interactions in local galaxies (Kennicutt 1998), and high-resolution *N*-body simulations including gas dynamics (Mihos & Hernquist 1994, 1995, 1996; Barnes & Hernquist 1996) have shown that collisions can trigger a bar-like instability that efficiently drives gas into centrally concentrated, high-density knots, creating the conditions that are likely to result in a starburst.

Our terminology implies a dichotomy between 'quiescent' and 'bursting' star formation that may be somewhat artificial – the recent work of Kennicutt (1998) suggests that star formation in both quiescent (normal) and starburst galaxies has the same dependence on the *local* gas density. By funnelling gas into a small central region, the interaction may simply produce the elevated gas densities that in turn lead to enhanced levels of star formation. Still, in interpreting the high-redshift observations it is important conceptually to determine what kind of process actually dominates early star formation; the difference between internally governed versus externally triggered star formation is highly significant. For example, it determines whether we interpret the number and luminosity density of high-redshift galaxies as reflecting basically their masses and internal properties, or instead the merger rate at that epoch and the efficiency of the gas inflows produced by these mergers.

In this paper we investigate a scenario in which most of the observable LBGs are collisional starbursts. Despite the fact that such collisions inevitably arise in hierarchical clustering models (e.g. Kolatt et al., in preparation), small-object collisions would not have been resolved in existing hydrodynamical simulations (e.g. Katz, Hernquist & Weinberg 1999), while previous semi-analytic work may have also underestimated the importance of collisional starbursts because of their assumed physical recipes. We investigate whether the numbers and properties of high-redshift starbursts are compatible with observations of LBGs, and whether this scenario leads to self-consistent agreement with global quantities such as the star formation rate density and the cold gas and metal content of the Universe as a function of redshift. We also consider models containing only 'quiescent' star formation, and discuss whether the collisional starburst picture is merely *compatible* with the data or whether some contribution from bursts seems to be *required*. Our conclusion is that quiescent models can marginally match all data provided that early star formation is highly accelerated, but that burst models fit all data naturally and are therefore preferred. Either way, we conclude that early star formation must be much more efficient than locally.

The outline of this paper is as follows. In Section 2 we give a brief introduction to the semi-analytic models, including our simple approach for including collisional starbursts. In Section 3 we present the predictions of our fiducial models and compare

them with the observations. In Section 4 we compare our results with previous work, and in Section 5 we discuss the sensitivity of our results to various assumptions, including the cosmology, stellar population synthesis and dust modelling. We summarize our results and conclude in Section 6. The non-specialist reader may wish to skip Sections 4 and 5, which are somewhat technical. The main results of these sections are summarized in a general way in Section 6.

2 SEMI-ANALYTIC MODELLING

2.1 Basics

We use semi-analytic techniques to model the formation and evolution of galaxies in a hierarchical clustering framework. Our models include the effects of gravitation on the formation and merging of dark matter haloes, the hydrodynamics of cooling, star formation, supernovae feedback and metal production, galaxy–galaxy merging, and the evolution of stellar populations. Our Monte Carlo-based approach is similar in spirit to the models originally presented by Kauffmann, White & Guiderdoni (1993, hereafter KWG93) and Cole et al. (1994, hereafter CAFNZ), and subsequently developed by these groups (hereafter referred to as the 'Munich' and 'Durham' groups) and others in numerous papers. The semi-analytic models used here are described in Somerville (1997) and Somerville & Primack (1999, hereafter SP). As shown in SP, these models produce good agreement with a broad range of local galaxy observations, including the Tully–Fisher relation, the *B*- and *K*-band luminosity functions, cold gas contents, metallicities, and colours. We refer the reader to SP for a more comprehensive review of the literature and a more detailed description of our models. Below we sketch our approach briefly.

The framework of the semi-analytic approach is the 'merging history' of a dark matter halo of a given mass, identified at $z = 0$ or any other redshift of interest. We construct Monte Carlo realizations of 'merger trees' using the method of Somerville & Kolatt (1999), which was tested against merger trees extracted from dissipationless simulations (Somerville et al. 2000). Each branch in the tree represents a halo merging event, and the trees are truncated when the circular velocity of the progenitor halo becomes smaller than $40\,\mathrm{km\,s^{-1}}$. We assume that gas in haloes smaller than this effective mass resolution is photoionized and cannot cool or form stars (see SP). We construct merger histories for a grid of haloes with circular velocities ranging from 40 to $1500\,\mathrm{km\,s^{-1}}$, and weight the results using the appropriate Press–Schechter probability for the appropriate halo mass and redshift.

When a halo collapses or merges with a larger halo, we assume that the associated gas is shock-heated to the virial temperature of the halo. This gas then radiates energy and cools. The cooling rate depends on the density, metallicity, and temperature of the gas. Cold gas is turned into stars using a simple recipe, and supernovae energy reheats the cold gas according to another recipe. To model the star formation rate, we use an expression of the general form

$$\dot{m}_* = \frac{m_{\rm cold}}{\tau_*}, \tag{1}$$

where $\dot{m}_*$ is the star formation rate, $m_{\rm cold}$ is the total mass of cold gas in the disc, and cold gas is converted into stars on a time-scale τ_*. In principle, this time-scale could be a function of the circular velocity or dynamical time of the disc, or other variables. In the simplest version of this recipe, $\tau_* = \tau_*^0 =$ constant; i.e., the efficiency of conversion of cold gas into stars $\dot{m}_*/m_{\rm cold}$ is independent of the other properties of the galaxy. This recipe was one of those investigated in SP (called SFR-C in the terminology of SP; here called 'constant efficiency quiescent'). We also consider a recipe in which $\tau_* \propto t_{\rm dyn}$, where $t_{\rm dyn}$ is the dynamical time of the galaxy. Because the density of the collapsed haloes is higher at high redshift, the typical dynamical times become smaller and the conversion of gas into stars becomes faster. This recipe is similar to the one usually used by the Munich group (called SFR-M in SP), and here we call it 'accelerated quiescent.'

Supernovae feedback is modelled using the disc–halo model introduced in SP. The rate of reheating of cold gas is given by

$$\dot{m}_{\rm rh} = \frac{\dot{\epsilon}_{\rm SN}}{v_{\rm esc}^2}, \tag{2}$$

where $\dot{\epsilon}_{\rm SN}$ is the rate at which energy is injected into the cold gas by supernovae, and $v_{\rm esc}^2$ is the average escape velocity of the disc or halo (the reheating rate and ejected gas mass is calculated separately for each of these components).

Each new generation of stars produces a fixed yield of metals, which are immediately deposited back into in the cold gas. The metals may then be mixed with the hot gas in the halo, or ejected from the halo by supernovae feedback. This in turn affects the cooling, which becomes more efficient as the metallicity of the gas increases. Unlike our work in SP, where we used a fixed metallicity for the hot gas in calculating the cooling function, here we use the modelled value of the hot gas metallicity and the metallicity-dependent radiative cooling curves tabulated by Sutherland & Dopita (1993).

Another new feature (not present in the models of SP) is the treatment of gas and metals that are ejected from the halo. In SP this material was never re-incorporated in any halo. Here we assume that the material is distributed outside of the halo with a continuation of the isothermal r^{-2} profile that we assume inside the halo, and such that if the total mass of the halo were to double, all of the material would be re-incorporated in the halo. This material falls in gradually (according to a spherical infall model) as the virial radius of the halo increases due to the falling background density of the Universe. This is similar to the recipe used by CAFNZ and Kauffmann et al. (1999), in which all of the hot gas ejected from the halo is re-incorporated (all at once) when the halo doubles in mass, but it removes the discontinuous behaviour caused by the sudden infall of a large amount of gas. The assumption that the ejected gas should return when the halo mass doubles is just an arbitrary relic of the block model used by CAFNZ, in which the halo masses always grew by factors of 2 in each branch of the merging tree; the actual time-scale for return of the ejected gas is quite uncertain.

We now also account for the mass that is returned to the ISM by young stars. For each quantity of mass that is turned into stars, a fraction R is returned to the cold gas reservoir. Here we assume $R = 0.14$, which results from assuming that all mass from stars with masses greater than $8\,\mathrm{M_\odot}$ is recycled, using a Salpeter IMF with an upper mass cut-off of $100\,\mathrm{M_\odot}$. Had we also included mass-loss from lower mass stars, the value of R would be larger, but these estimates are more uncertain, so we neglect this contribution.

2.2 Mergers and starbursts

When dark matter haloes merge, the galaxies contained within them survive and may merge on a longer time-scale. After a halo merger event, the central galaxy of the largest progenitor

halo becomes the new central galaxy, and all other galaxies become 'satellites.' Satellite galaxies may merge with the central galaxy on a dynamical friction time-scale, or with each other on approximately a mean free path time-scale. However, if the relative velocities of the satellites are large compared to their internal velocities, they will not experience a binding merger. Thus the satellite merger rate decreases in clusters. The expressions for the satellite–central and satellite–satellite merger rates are given in SP.

We assume that every galaxy–galaxy merger triggers a starburst. Our treatment is based on a simple parametrization of the results of *N*-body simulations with gas dynamics, which we now summarize briefly. Mihos & Hernquist (1994, 1996) have simulated galaxy–galaxy mergers using a high-resolution *N*-body/smoothed particle hydrodynamics code (TREESPH) with star formation modelled according to a Schmidt law ($\rho_{\rm SFR} \propto \rho_{\rm gas}^{n}$ with $n = 1.5$). In Mihos & Hernquist (1996), mergers between equal-mass galaxies (major mergers) were simulated, and it was found that 65–85 per cent of the total gas supply (in both galaxies) was converted into stars over a time-scale of 50–150 Myr. These results were not very sensitive to morphology or the orbital geometry.

The case of highly unequal-mass mergers was explored in Mihos & Hernquist (1994). The case simulated represents a Milky Way-sized disc galaxy accreting a satellite that is one-tenth of its mass. The non-axisymmetric mode generated by the accretion of the satellite causes a large fraction of the gas to collapse into the central region of the galaxy, fuelling a strong starburst. In this case about 50 per cent of the original gas supply is consumed in the starburst, which lasts for about 60 Myr. However, the results are much more sensitive to the morphological structure of the galaxies than the major-merger case. If the larger galaxy has a bulge (the case simulated has a bulge to disc mass ratio of 1:3), the bulge seems to stabilize the disc against strong radial gas flows, leading to a much weaker starburst event (only about 5 per cent of the total gas supply is consumed). Mihos & Hernquist (1994) note that this implies that the importance of bursts will decrease at low redshift as galaxies develop bulges, even if the merger rate remains constant.

These results inform our treatment of collisional starbursts. When any two galaxies merge, the 'burst' mode of star formation is turned on. The star formation rate due to the burst is modelled as a Gaussian function of time with a width of $\sigma_{\rm burst}$. The burst is switched off after a time $4\sigma_{\rm burst}$ has elapsed. The burst model has two parameters, the time-scale and the efficiency of the burst. The efficiency $e_{\rm burst}$ is defined as the fraction of the cold gas reservoir (of both galaxies combined) that is turned into stars over the entire duration of the burst. We assume that the efficiency is a power-law function of the mass ratio of the merging galaxy pair:

$$e_{\rm burst} = f_{\rm consume}(m_{\rm small}/m_{\rm big})^{\alpha_{\rm burst}}, \qquad (3)$$

with the two parameters $f_{\rm consume}$ and $\alpha_{\rm burst}$ chosen to reproduce the results of Mihos & Hernquist (1996) for a 1:1 merger, and Mihos & Hernquist (1994) for a 1:10 merger ($f_{\rm consume} = 0.75$ and $\alpha_{\rm burst} = 0.18$). When a bulge of at least a third of the disc mass is present, bursts are suppressed in minor mergers (based on Mihos & Hernquist 1994). Fig. 1 shows the resulting scaling of the burst efficiency with the mass ratio of the merging pair for the bulge and no-bulge cases. We assume that the burst time-scale $\sigma_{\rm burst}$ is proportional to the dynamical time of the larger of the two discs. The quiescent mode of star formation continues uninterrupted according to equation (1) above, although its contribution is generally insignificant in comparison to the burst.

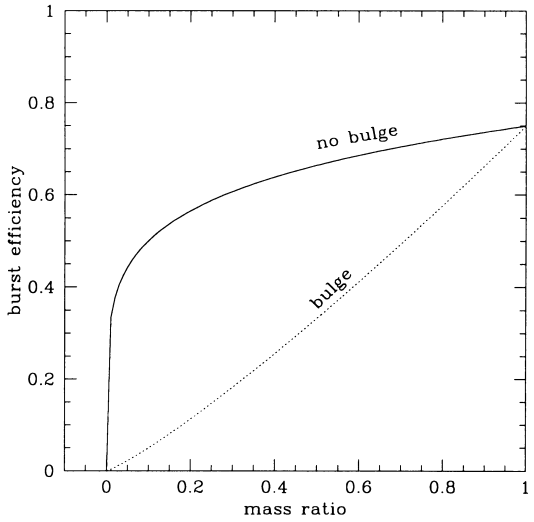


Figure 1. The efficiency of star formation in bursts as a function of the baryonic mass ratio of the merging galaxies. The results are scaled to match the calculations of Mihos & Hernquist (1994, 1996). When a bulge is present, the burst efficiency drops off more quickly at lower mass ratios.

Each merger is classified as 'major' or 'minor' according to whether the ratio of the smaller to the larger of the galaxies' baryonic masses is greater than or less than the value of the parameter $f_{\rm bulge} \sim 0.25$. Major mergers have mass ratios greater than $f_{\rm bulge}$, and the bulge and disc stars of both galaxies plus all new stars formed in the burst are placed in a bulge component. Minor mergers have mass ratios less than $f_{\rm bulge}$, and the pre-existing stars from the smaller galaxy are placed in the disc component of the post-merger galaxy, while all newly formed stars again join the bulge. This is motivated by the *N*-body simulations of similar satellite accretion events by Walker, Mihos & Hernquist (1996), which show that 90 per cent of the satellite mass is stripped and distributed in the disc of the larger galaxy, and only the densest part of the satellite core ends up at the centre.

A possible problem in scaling from the $z = 0$ simulations is that galactic discs may be quite different at high redshift. Galaxies are then smaller and denser, and probably have higher ratios of gas to stars. This may significantly affect their susceptibility to star-forming instabilities. We are currently investigating this using a more extensive set of *N*-body hydrodynamic simulations similar to those of Mihos & Hernquist, but with the initial galaxy properties chosen to be representative of high-redshift galaxies (Somerville et al., in preparation). Similarly, the modelling of mergers, infall and tidal stripping of subhaloes should be refined and tested against high-resolution dissipationless cosmological simulations. Work on these issues is also in progress (Kolatt et al. 2000).

Fig. 2 shows the total star formation rate for the largest progenitor of the central galaxy in a Milky Way-sized halo ($V_{\rm c} = 220\,{\rm km\,s^{-1}}$) and in a group-sized halo ($V_{\rm c} = 500\,{\rm km\,s^{-1}}$), both at $z = 0$ today. The star formation rate is shown in models with (1) no starbursts, (2) bursts in major mergers only, and (3) bursts in both major and minor mergers. All three models contain quiescent star formation using the constant efficiency recipe. From this figure we can see that minor mergers are important even if their efficiency is low, because they are much more common than major

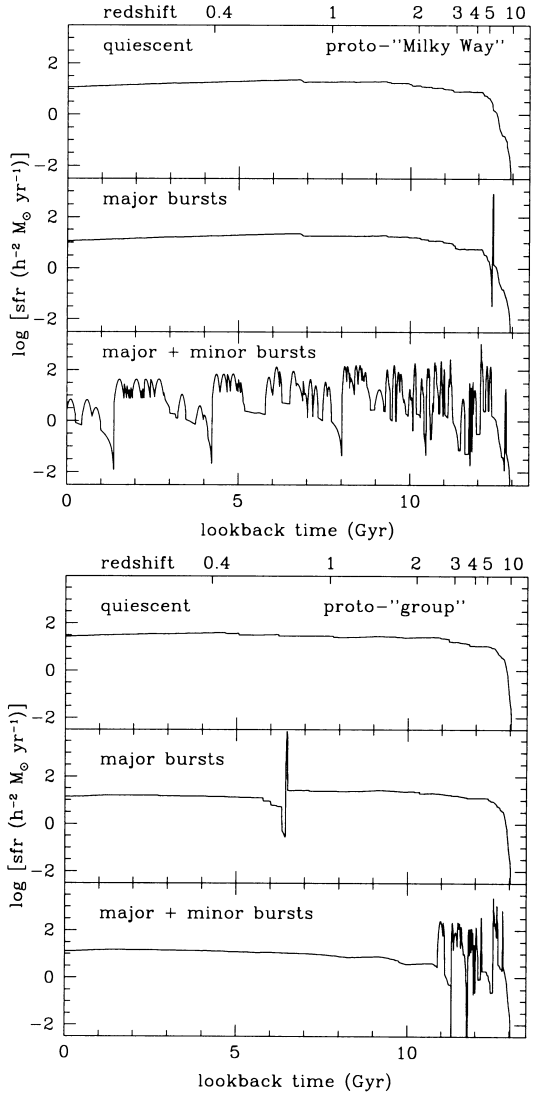


Figure 2. The total star formation rate for the largest progenitor of the central galaxy within a halo with a present-day circular velocity of 220 or 500 km s^{-1}. All three models contain constant efficiency quiescent star formation. The top panel shows the model with no bursts, the middle panel shows the model with bursts in major mergers only, and the bottom panel shows the models with bursts in major and minor mergers.

mergers. Comparing the group-sized halo with the Milky Way halo shows how mergers shut down in high-velocity-dispersion environments. This is because we have assumed that the relative velocity of colliding objects must be small compared to their internal velocities if objects are to merge. The burst models discussed in the remainder of this paper correspond to model (3) unless specified otherwise.

2.3 Stellar population synthesis

With the full star formation history of a galaxy in place, we estimate its luminosity in any desired passband using the stellar population synthesis models of Bruzual & Charlot (1993) (GISSEL). In this paper our standard treatment uses the 1998 version of the solar-metallicity models with a Salpeter IMF. Modelling the young, massive, subsolar metallicity stars that might be found in LBGs is probably quite uncertain (see Charlot, Worthey & Bressan 1996). However, it is encouraging that several groups have produced models that agree quite well in the UV and optical part of the spectral energy distribution (SED) (see Devriendt, Guiderdoni & Sadat 1999 and Section 5.2). The mass-to-light ratio in the UV is relatively insensitive to metallicity for young stars, but it is quite sensitive to the IMF (primarily to the ratio of high-mass to low-mass stars). We include only starlight and neglect the contribution from nebular emission. See Section 5.2 for a detailed discussion of uncertainties in the stellar population synthesis.

2.4 Dust

Because most of the observations of high-redshift galaxies are obtained in the rest-UV, the effects of dust are likely to be important. Just *how* important has been a matter of debate ever since the discovery of the LBG population. Pettini et al. (1997a) estimated an extinction at $\sim$1500 Å of a factor of $\sim$3, based on the ratio of emission lines to the continuum in a few objects. Meurer et al. (1997) and Sawicki & Yee (1998) estimated much larger factors of 15–20. Recently, there seems to have been convergence to an intermediate value of a factor of $\sim$5 (Meurer, Heckman & Calzetti 1999; Steidel et al. 1999). These authors made use of a correlation between the FIR excess (a reliable observational measure of bolometric extinction) and the far-UV spectral slope in nearby starburst galaxies. If the same correlation is then assumed to hold in high-redshift galaxies, the measured UV spectral slopes of LBGs indicate that there is an average extinction of a factor of 4.7 in the relatively bright ($\mathcal{R} < 25.5$) LBG population studied by Steidel et al. (1999). The same measurements (Meurer et al. 1999) indicate that, in LBGs as in local starburst galaxies, the most UV-luminous (and rapidly star-forming) objects are the most heavily extinguished (Wang & Heckman 1996).

To estimate the effects of dust on our model galaxies, we use a very simple parametrization of the empirical results of Wang & Heckman (1996) for nearby starburst galaxies. The optical depth for a galaxy with an intrinsic (unextinguished) UV luminosity $L_{\rm UV,i}$ is given by

$$\tau_{\rm UV} = \tau_{\rm UV,*}\left(\frac{L_{\rm UV,i}}{L_{\rm UV,*}}\right)^{\beta}. \quad (4)$$

This is identical to the recipe used in SP, except that we now normalize the recipe in the UV rather than the B band. We take $L_{\rm UV,*}$ to equal the observed value of L_* in the $z = 3$ sample of Steidel et al. (1999), i.e., the luminosity corresponding to $m_{AB} = 24.48$ in the appropriate cosmology. We use the same value of L_* at redshift $z = 4$, which, as Steidel et al. (1999) note, is consistent with the $z = 4$ data. We take $\beta = 0.5$ as in Wang & Heckman (1996). We then assign to each galaxy a random inclination, and compute the actual extinction using a standard slab model (see SP). We adjust the parameter controlling the face-on optical depth, $\tau_{\rm UV,*}$, to obtain an average extinction correction at 1500 Å of a factor of $\sim$5 for $z \sim 3$ galaxies with luminosities typical of the Steidel et al. sample, consistent with the estimates described above. The value $\tau_{\rm UV,*} = 1.75$ gives good results. To extend the extinction correction to other wavebands, we assume a Calzetti attenuation curve (Calzetti 1997).

2.5 Model parameters

Throughout this paper, unless stated otherwise, we use the

fashionable ΛCDM cosmology with $\Omega_0 = 0.3$, $\Omega_\Lambda = 0.7$, $h \equiv H_0/(100\,\mathrm{km\,s^{-1}\,Mpc^{-1}}) = 0.7$, $\sigma_8 = 1.0$, and a baryon density of $\Omega_b = 0.019\,h^{-2}$. These values are consistent with a great deal of current data (for a recent review see Primack 2000). As in SP, the main free parameters related to galaxy formation – those describing the quiescent star formation efficiency (τ^0_*), the supernovae feedback efficiency ($\epsilon^0_{\rm SN}$), and the metallicity yield (y) – are set by requiring an average present-day 'reference galaxy' (the central galaxy in a halo with a circular velocity of $220\,\mathrm{km\,s^{-1}}$) to have (a) an I-band magnitude $M_I - 5\log h = -21.8$, (b) a cold gas mass $m_{\rm cold} \simeq 10^{10}\,h^{-2}\,\mathrm{M_\odot}$, and (c) a stellar metallicity of about solar. Requirement (a) adjusts the zero-point of the I-band Tully–Fisher relation to agree with observations, while requirement (b) fixes the cold gas content of the 'reference galaxy' to agree with observed H I masses, allowing for a contribution from molecular hydrogen (see SP). The values of the free parameters used here are similar to those used in SP.

In some previous work (e.g. BCFL), it has been assumed that only a fraction $f^*_{\rm lum} < 1$ of the stars formed in the models are luminous, the remainder being in the form of brown dwarfs or other non-luminous baryonic material. We find that in order to get our reference galaxy to be bright enough today with our assumed Salpeter IMF, we need $f^*_{\rm lum} \simeq 1$.

Note that the parameters have been set entirely by comparing to a subset of present-epoch observations. We have shown in SP that the models then reproduce many important observed features of nearby galaxies. Now, leaving all free parameters fixed, we can consider the predictions of the same models for the high-redshift Universe.

3 MODEL RESULTS

Because extensive spectroscopic follow-up work has now verified the effectiveness of the Lyman-break or 'drop-out' technique for finding $z \gtrsim 2$ galaxies, we do not attempt to mimic the colour selection process used to identify real Lyman-break objects, but rather assume that all galaxies brighter than a certain limiting magnitude in our models would in fact be selected as LBGs. The modelling work of BCFL has shown that using the same colour selection criteria as the observations picks out model galaxies in the expected redshift range. We do not include the effect of absorption by intervening cold gas clouds because, although this effect is very important shortwards of the Lyman limit, for redshifts less than $z \sim 4$ it does not much affect the SED in the wavelength range relevant to our study. We have calculated magnitudes using the filter response functions of the WFPC2 filters F606W and F814W (V_{606} and I_{814}), as well as the $\mathcal{R}$ and I filters of Steidel & Hamilton (1992) used in the ground-based observations and kindly provided to us in electronic form by K. Adelberger. All magnitudes are given in the AB system (Oke & Gunn 1983). The V_{606} filter and the $\mathcal{R}$ filter correspond to a mean rest wavelength of 1600 Å at redshifts 2.75 and 3.04 respectively. The I_{814} and Steidel & Hamilton I filters correspond to the same rest wavelength at redshifts of about 4.00 and 4.13 respectively.

We now consider three different recipes for star formation. The model referred to as the 'collisional starburst' model includes quiescent star formation using the 'constant efficiency' (CE) recipe, plus star formation in bursts as described in the previous section. The 'constant efficiency quiescent' model has no burst mode, and all star formation is modelled using the CE recipe. The third model is also quiescent, but star formation efficiency varies inversely as the dynamical time of the galaxy. We refer to this third model as 'accelerated quiescent', because the star formation rate for a given cold gas mass becomes *higher* at high redshift due to the increasing density of the haloes. Note that the supernovae feedback recipe has been kept the same, and all other free parameters have been left fixed to the same values for the three models. In SP we showed that, by adjusting only the normalization of the star formation efficiency, τ^0_*, it is possible to make all three models match observed properties of local galaxies such as the B- and K-band luminosity functions, the Tully–Fisher relation, colours, gas and metal contents, etc.

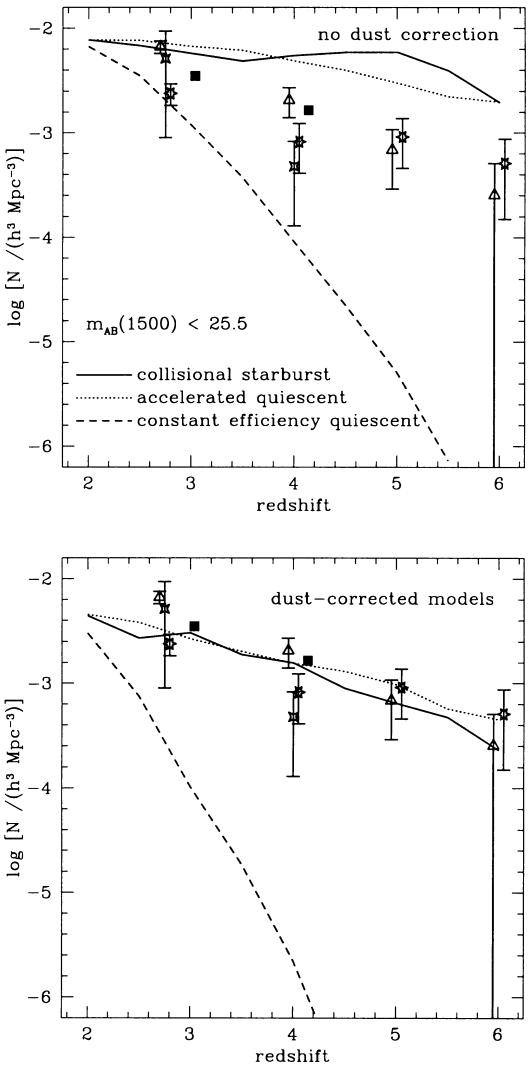


Figure 3. The comoving number density of galaxies brighter than 25.5 at rest 1500 Å. The upper panel shows the three models with no correction for dust extinction, and the lower panel shows the models with dust extinction included using the recipe described in Section 2.4. The filled squares indicate the comoving number density of LBGs with spectroscopic redshifts from the ground-based sample of Steidel et al. (1999). Four-pointed stars indicate the values derived from LBGs in the HDF (Pozzetti et al. 1998). The stars and triangles are from the HDF-N and HDF-S galaxies with photometric redshifts from the Stonybrook catalogues (Lanzetta et al. 1999; Chen et al. 1999a). The observations have not been corrected for extinction. Note that the CE quiescent model predicts a strong decrease in the number density of bright galaxies with redshift, in contrast to the other two models, which show a very gentle decline from $z \sim 2$ to $z \sim 6$.

3.1 Comoving number density

We first investigate the model predictions for the number density of bright objects as a function of redshift. Fig. 3 shows the comoving number density of galaxies brighter than an apparent AB magnitude limit of $m_{AB}(1500) = 25.5$ at rest wavelength 1500 Å over the redshift range $2 \leq z \leq 6$. The comoving number densities for the observations have been calculated using the appropriate geometry for the ΛCDM cosmology used in our models. For the ground-based sample of LBGs with spectroscopic redshifts, we have used the values given in table 3 of Steidel et al. (1999). The 1σ field-to-field variance of this sample is about 12 per cent; if plotted on Fig. 3, the error bar would be smaller than the symbol. The comoving number density of LBGs in the HDF has been calculated using the V_{606} counts of U-dropouts from table 1 of Pozzetti et al. (1998), and the I_{814} counts of B-dropouts helpfully provided to us by L. Pozzetti. Error bars on these points represent Poisson errors only. We also show the number densities computed from the Stonybrook catalogue of photometric redshifts from the HDF-N and HDF-S (Lanzetta et al. 1999; Chen et al. 1999a), generously provided to us by H.-W. Chen and K. Lanzetta. Once again the error bars are simply 1σ Poisson and do not account for photometric or redshift errors. Note that, unlike the point derived from galaxies identified in the HDF using the Lyman-break technique (Pozzetti et al. 1998), the values from the photometric redshift catalogues at $z \sim 3$ agree with the ground-based estimate within the errors. This suggests that the full photometric redshift technique, including the observed near-IR bands, finds high-redshift galaxies that are missed by the Lyman-break technique. The difference between the HDF-N and HDF-S (stars and triangles) gives an indication of the field-to-field variance characteristic of the HDF volume.

Fig. 3 demonstrates an important and robust prediction of the models – the *constant efficiency* quiescent models predict a *strong decline* in the number density of bright galaxies with increasing redshift. This is a generic feature of any hierarchial model in which the efficiency of the conversion of cold gas into stars remains constant with redshift. This is because the number density of the massive haloes that host bright galaxies in the quiescent model drops off sharply with redshift. Kolatt et al. (1999) found a similar result based on *N*-body simulations. The net result is that the CE quiescent model without dust extinction produces an acceptable match to the number of bright galaxies at $z \sim 3$, but badly underpredicts the numbers at redshifts $z \geq 4$. The shortfall is unacceptable at all redshifts when dust is included.

In contrast, the burst model contains collisional starbursts that cause lower mass objects to become bright enough to exceed the detection limit for a brief amount of time. Because the collision rate is relatively constant over the redshift range $2 \lesssim z \lesssim 6$ and there is an ample supply of cold gas, the comoving number density of visible galaxies remains *almost constant* over this range. The collisional starburst model actually overpredicts the number of bright galaxies, but adding dust provides excellent agreement at all redshifts. Finally, the accelerated quiescent model produces nearly identical results to the collisional starburst model, but for a very different reason. Instead of being temporarily brightened by bursts, each galaxy has a constant low mass-to-light ratio because of the accelerated star formation rate. This model also does a good job of matching the number of moderately bright galaxies at all redshifts.

The observations at $2 \lesssim z \lesssim 4.5$, especially those based on the large ground-based sample of Steidel et al. (1999), provide a secure lower limit on the number density of bright galaxies. The situation at higher redshift is more uncertain, but the differences between the CE quiescent model and the other two are quite dramatic. According to the CE quiescent model, the probability of finding even one bright galaxy at $z \gtrsim 5$ in a volume comparable to the HDF is only one in 10^6, whereas both the collisional starburst model and the accelerated quiescent model predict on the order of a few galaxies per HDF at $4.5 \lesssim z \lesssim 6$, even with dust. A number of candidate high-redshift galaxies with $z \gtrsim 5$ have been discovered recently (Dey et al. 1998; Spinrad et al. 1998; Weymann et al. 1998; Chen, Lanzetta & Pascarelle 1999b; Hu, McMahon & Cowie 1999; van Breugel et al. 1999), but, as most of these objects were found serendipitously, it is difficult to estimate their number density. Such candidates are also being detected in fair numbers in photometric redshift catalogues, but these detections are also uncertain, and further spectroscopic work will be required in order to confirm them (see Stern & Spinrad 2000 for a recent review on search techniques for very high-redshift galaxies). If the current preliminary detections of very high-redshift galaxies are confirmed, they will essentially rule out the CE quiescent model.

3.2 Luminosity function

The luminosity function and its evolution with redshift provide additional tests of models. The bright end ($m_{AB} \lesssim 25.5$) of the luminosity function at ∼1500 Å for the observed LBGs at $z \sim 3$ can be determined fairly accurately from the relatively large ground-based spectroscopic sample. The HDF can then be used to give some indication of the faint-end slope, although there is probably some incompleteness faintwards of about $m = 26$ (Dickinson 1998). The resultant composite luminosity function at $z = 3$ is shown in Figs 4 and 5. The ground-based data have been corrected for incompleteness (see Steidel et al. 1999). As

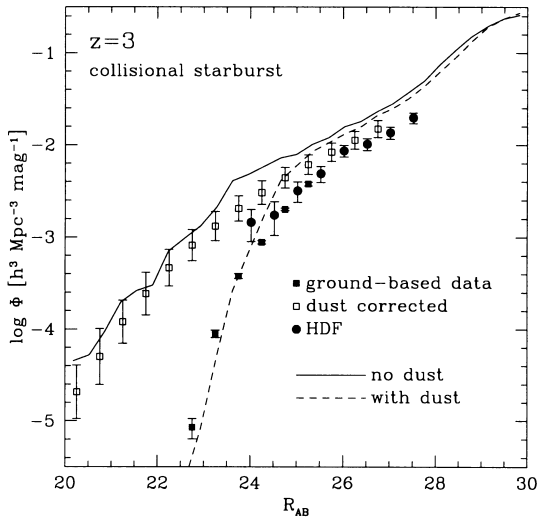


Figure 4. The luminosity function at $z = 3$ in the observed $\mathcal{R}_{AB}$ band (rest ∼ 1600 Å). Symbols show the composite luminosity function from ground-based observations (filled squares; Steidel et al. 1999) and the HDF (filled circles; Pozzetti et al. 1998). Open squares show the luminosity function derived from the ground-based observations after correcting for the effects of dust extinction using the method described in the text (Adelberger & Steidel 2000). Lines indicate the results of the collisional starburst model without (solid) and with (dashed) corrections for dust extinction.

usual, the observations have been converted to our ΛCDM cosmology.

Fig. 4 shows that the luminosity function of the collisional starburst model with dust has a steep, Schechter-type drop-off at bright magnitudes, whereas the intrinsic function without dust is much flatter. Thus, in this picture the knee seen in the observations is purely an artefact of differential dust extinction. It is gratifying that our intrinsic luminosity function is very similar in shape to the observed dust-corrected function derived from the data (Adelberger & Steidel 2000), for which the corrected luminosity of each galaxy has been derived using the measured far-UV slope and the dust–slope correlation discussed in Meurer et al. (1999).

Fig. 5 shows similar plots for the two quiescent models. The CE model without dust (bottom panel) has a steeply decreasing luminosity function at the bright end, reflecting the steepness of the halo mass function at these large masses. When dust extinction is added, the CE quiescent model becomes even steeper and dramatically underpredicts the number of bright galaxies, consistent with its underprediction of count data in Fig. 3. The situation with the accelerated quiescent model is rather different. Although that model predicts about the same total number of bright [$m_{AB}(1500) < 25.5$] galaxies as the collisional starburst model in Fig. 3, the *shape* of its luminosity function is different at the very brightest magnitudes. The burst model shows a tail at bright magnitudes, whereas the accelerated quiescent model, even *without* dust, drops off steeply, like a Schechter function. With dust included, the accelerated quiescent model falls even more steeply and fails to match the brightest galaxies. This, again, is due to the tight link between halo masses and galaxy luminosities in all quiescent models.

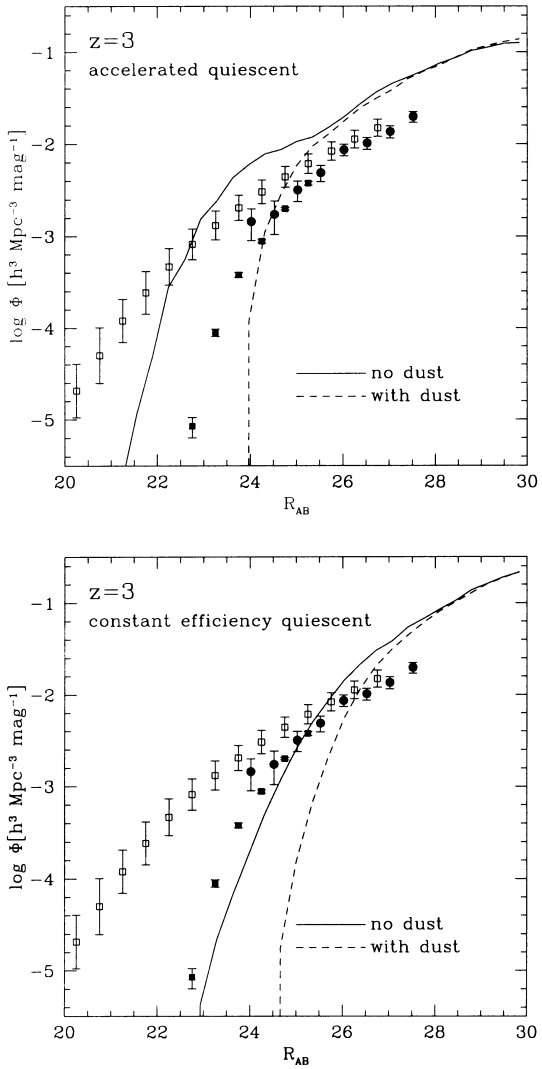


Figure 5. Same as Fig. 4, for the quiescent models.

Figs 6 and 7 show analogous comparisons for luminosity functions at $z \sim 4$. The collisional starburst model fits the luminosity function for bright galaxies, whereas both quiescent models fall short, especially when dust is added. All models predict too many faint galaxies. This may be due to the fact that, with the parameters we have chosen, our dust model predicts very little extinction in galaxies fainter than $m \simeq 26$. In correcting the models for dust, we have simply assumed the same parameters as at $z \sim 3$. Note also that the observed luminosity function faintwards of $m \simeq 25$ (the magnitude reached by the ground-based survey) is quite uncertain and possibly incomplete. Both effects would tend to make the models deviate in this way.

In summary, the CE quiescent models predict a strong evolution of L_* with redshift, whereas the burst model and the accelerated quiescent model predict very little evolution in L_* over the redshift range $3 \lesssim z \lesssim 5$. The observed and dust-corrected luminosity functions of the collisional starburst model are in good agreement with all data at both redshifts. However, systematic errors in the dust correction could be large. A more direct test of the number densities of very bright galaxies (which correspond to objects with extremely high star formation rates) will come from future submillimetre observations.

3.3 The star formation history of the Universe

The star formation history of the Universe, i.e., the global rate of star formation as a function of redshift, is a crucial test of theories of galaxy formation and cosmology. There are many observational tracers of the star formation history, each with their own selection

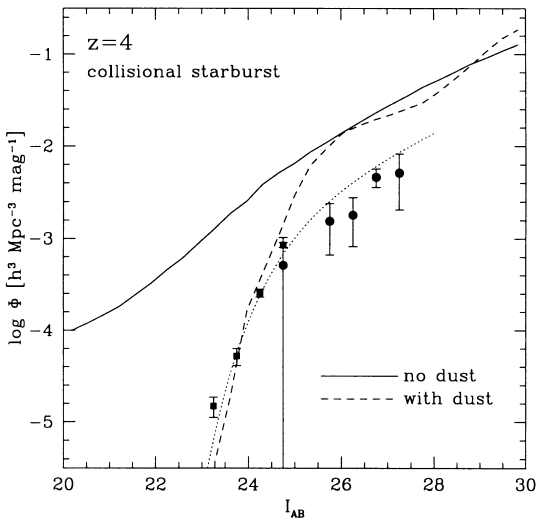


Figure 6. Same as Fig. 4, for the I band at $z = 4$ (rest ~ 1600 Å). Observed data are shown as solid symbols (see key in Fig. 4). The dotted line shows the Schechter fit to the $z = 4$ luminosity function of Steidel et al. (1999), which is somewhat higher than the HDF results. The data may be affected by incompleteness below ∼26 mag.

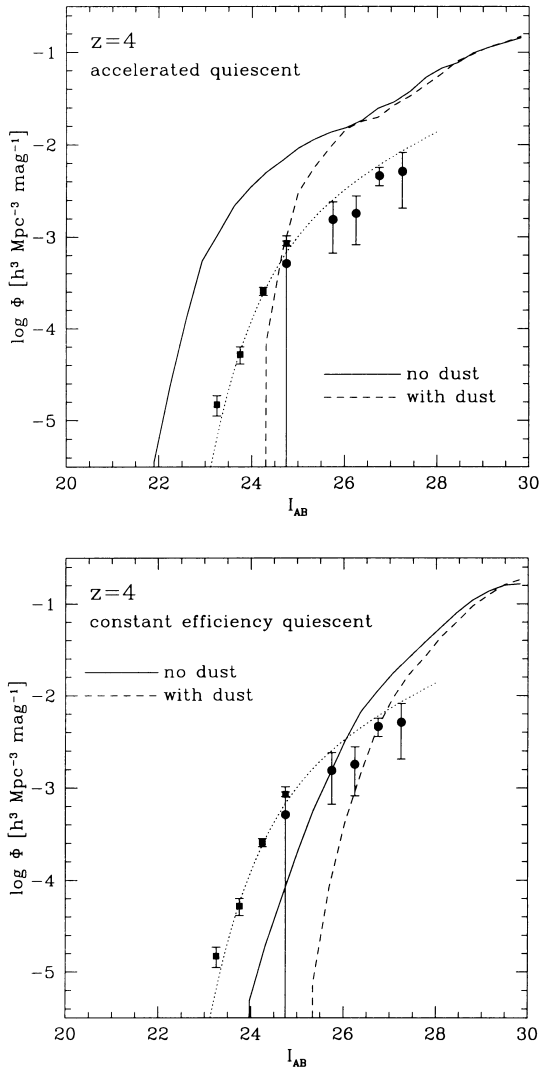


Figure 7. Same as Fig. 6, for the quiescent models.

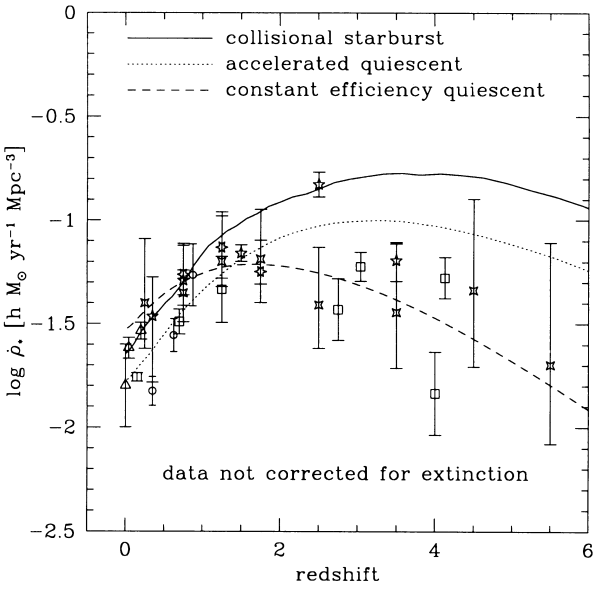


Figure 8. The star formation history of the Universe ('Madau diagram') for our ΛCDM cosmology. Symbols indicate observed estimates of the star formation rate density at various redshifts; sources of data are listed in Table A2. All data points have been corrected for incompleteness but *not* for dust extinction. Lines show predictions of the three models.

effects. In this section we investigate the star formation history traced by the luminosity density, age constraints on stars in nearby galaxies, cold gas detected in quasar absorption systems, and metals.

3.3.1 Luminosity density

The star formation history of the Universe, as represented by the popular 'Madau diagram' (Madau et al. 1996), is an important signature of any scenario of galaxy formation. However, an accurate determination of this diagram is far from straightforward: the observations must be converted from luminosity to star formation rate and corrected for incompleteness and dust extinction (see Appendix A). As new observations have been added at various redshifts, the diagram has undergone a continuing metamorphosis from its original form. Fig. 8 shows the star formation history obtained from our burst and quiescent models, along with a compilation of recent observations, where the observations have been corrected for incompleteness but *not* for dust extinction. The observations have been converted to our ΛCDM cosmology as described in Appendix A. Note that the CE quiescent model shows the characteristic peak at $z \sim 1.5$ and the relatively steep decline at higher redshifts that were notable features of the original Madau diagram (Madau et al. 1996). However, the data points in that diagram at $z = 2.75$ and $z = 4$ were based on the very small HDF survey and relied on the Lyman-break technique to select galaxies in the desired redshift range. As Madau et al. correctly emphasized, these points should be regarded as lower limits. More recent observations based on much larger, ground-based surveys with spectroscopic redshifts (Steidel et al. 1999) show a much shallower evolution between $z \sim 3$ and $z \sim 4$.

Another important recent observational development is a more convincing estimate of the effects of dust extinction on the values derived from luminosities in the far-UV. This has also produced a significant modification of the diagram, as recently noted by (Steidel et al. 1999). Fig. 9 shows a subset of the observations with a correction for dust extinction based on the observational results (see Appendix A for details). The CE quiescent model is now seen to be highly inconsistent with the revised observational estimates at high redshift. The collisional starburst model shows a gentler rise and a broad peak around $z \sim 4$, with near-constant star formation density out to a redshift of $\sim$6. This appears very consistent with the recent optical estimates after correction for dust extinction (see Appendix A); it is also consistent with the luminosity densities implied by the far-IR and sub-mm sources detected at intermediate and high redshift (Hughes et al. 1998; Flores et al. 1999). The redshift dependence of the star formation rate density in the accelerated quiescent model is very similar to the collisional starburst model, although the overall star formation rate density is a bit lower than the data.

Fig. 10 shows the contribution to the star formation rate density from the burst and quiescent modes separately for our collisional starburst model. The bursting mode dominates at high redshift, but declines more steeply than the quiescent mode as redshift decreases. A cross-over therefore occurs, at $z \sim 0.8$, and quiescent star formation dominates the low-redshift Universe.

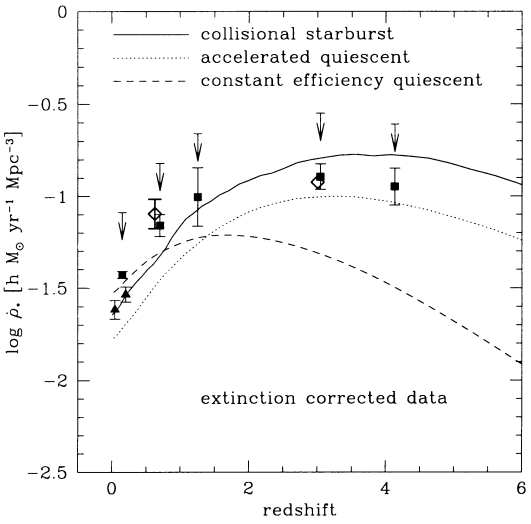


Figure 9. The new Madau diagram, corrected for dust extinction. Symbols indicate observational estimates of the star formation rate density at various redshifts. Triangles are derived from Hα (Gronwall 1998; Tresse & Maddox 1998), and squares are derived from rest-UV luminosities (Treyer et al. 1998; Cowie, Songaila & Barger 1999; Steidel et al. 1999). These data have been corrected for incompleteness and dust extinction (see Appendix A for details). The upper-limit symbols show the data with a 'maximal dust correction,' in which a fixed factor of 5 correction has been applied to all galaxies. If dust extinction is luminosity-dependent, this probably overestimates the true extinction correction (see text). Diamonds indicate results derived from far-IR sources at $z = 0.7$ observed by *ISO* (Flores et al. 1999) and sub-mm sources, believed to be at high redshift ($\sim$3), observed by SCUBA (Hughes et al. 1998). Lines show the three usual models.

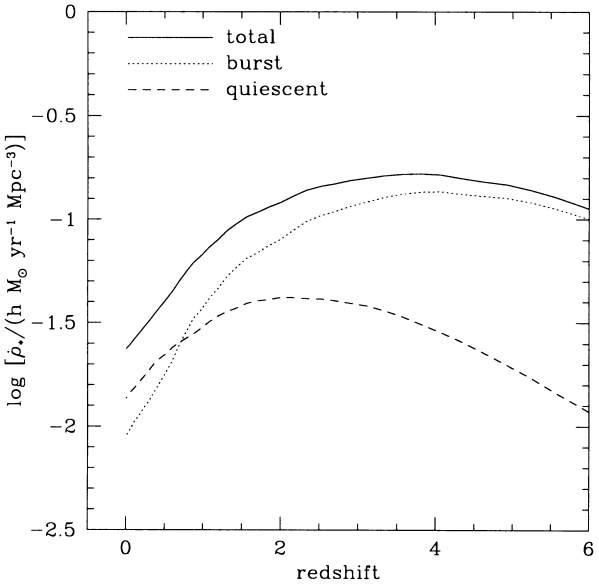


Figure 10. The star formation rate in the collisional starburst model is shown broken down into the contribution from quiescent star formation and bursts. The bursting mode dominates at high redshifts, and the quiescent mode dominates at low redshifts. The cross-over occurs at a redshift of about 0.8.

Fig. 11 shows the contribution to star formation in the burst mode from different kinds of merger events. We show the contribution from major and minor mergers between satellite and central galaxies, and major and minor mergers between two satellite galaxies. Note that in previous semi-analytic models, starbursts were included only in the first case, i.e., satellite–central/major mergers. We can see from the figure that the contribution from other kinds of mergers, particularly satellite–central/minor, is quite important. The relatively small contribution of satellite–satellite mergers may be an artefact of our rather questionable simplifying assumption that all new cold gas is accreted on to the central galaxy; this assumption tends to deprive small satellites of the fuel needed to produce bright starbursts, but may not be realistic.

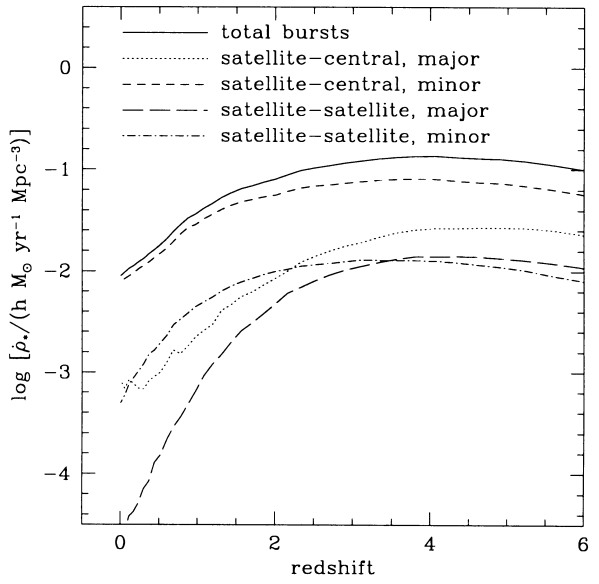


Figure 11. Star formation in the bursting mode is shown broken down into contributions from different kinds of merger events. The dotted and dashed lines show contributions from major and minor mergers (respectively) between a satellite galaxy and a central galaxy. The long-dashed and dot-dashed lines show the contribution from major and minor mergers between two satellite galaxies. The dominant contribution at all redshifts comes from minor mergers on to central galaxies.

3.3.2 Fossil evidence

Another way of constraining the star formation history of the Universe is from the 'fossil evidence' contained in the ages of stars in present-day galaxies. Renzini (1998) has argued that constraints on the ages of early-type galaxies in clusters from the small observed scatter of Fundamental Plane and colour–magnitude relations, combined with the fraction of today's stellar mass contained in early-type systems (bulges and ellipticals), can be used to deduce that one-third of the stars we see today must have formed at $z \gtrsim 3$. Van Dokkum et al. (1998) find that the formation redshift of early-type galaxies in clusters may be relaxed to $z \gtrsim 2$ in cosmologies with a large cosmological constant (such as our ΛCDM cosmology) because of the longer time that has ellapsed in such models. Fig. 12 shows the density of stars in units of the critical density for the burst and quiescent models. The point with the large error bar at redshift zero is from the 'baryon budget' of Fukugita, Hogan & Peebles (1998). All three models agree with this estimate within the errors. The shaded patch shows the constraints on Ω_* at high redshift from a Renzini-like argument, where we have simply divided the stellar

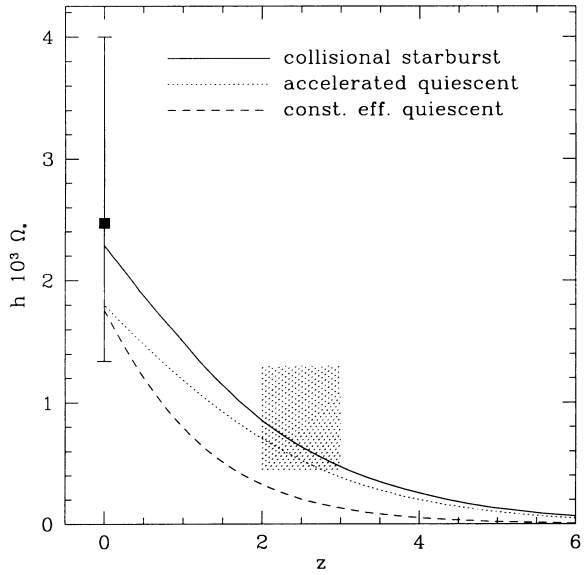


Figure 12. The density in stars, in units of the critical density (Ω_*). The observational estimate at $z = 0$ is shown by the symbol with the error bar. The observational constraint at high redshift from 'fossil evidence' (see text) is shown by the shaded box. The collisional starburst model and the accelerated quiescent model produce a significant fraction of their stars at high redshift, whereas the CE quiescent model produces stars too late.

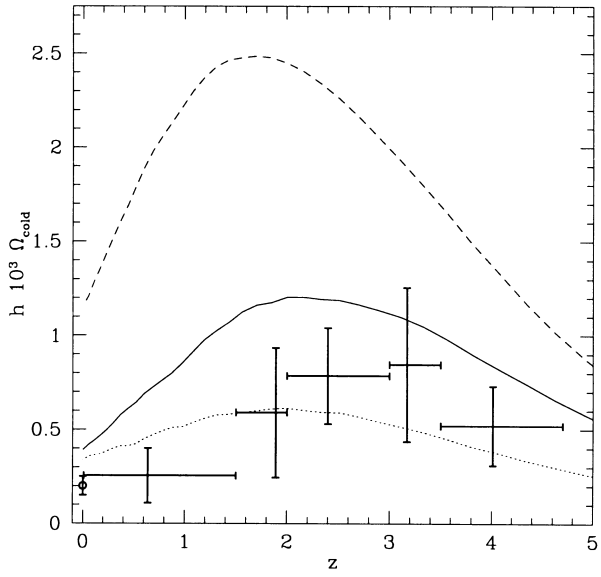


Figure 13. The evolution of $\Omega_{\rm cold}$ as a function of redshift. Data points show the density in the form of H I from observations of DLAS (Storrie-Lombardi, McMahon & Irwin 1996), converted to the ΛCDM cosmology. The data point at $z = 0$ is from local H I observations (Zwaan et al. 1997). The quantity plotted for the models (line types are as in Fig. 12) is generic 'cold gas', some of which may be in the form of molecular or ionized hydrogen, or may be in dust extinguished or low-column density systems that would not be detected as damped systems. Thus the model lines are upper limits on the quantity that is actually determined from the observations ($\Omega_{\rm HI}$ in DLAS) and should lie *above* the data points. Line types are the same as in Fig. 8.

baryon budget at $z = 0$ (and its error bar) by a factor of 3. The CE quiescent model does not produce enough stars at high redshift, but the collisional starburst and accelerated quiescent models meet this constraint. The CE quiescent model produces one-third of the stars by $z = 1.5$, and half of the stars by $z = 0.9$. The collisional starburst and accelerated quiescent models produce one-third of their stars by $z = 2.2$, and half of their stars by $z = 1.5$.

3.3.3 Cold gas density

Observations of quasar damped Lyman-α systems (DLAS) provide an estimate of the H I and metal content of the Universe from $z \sim 0.7$ to $z \sim 4$. Fig. 13 compares the models to an estimate of the fraction of the critical density in the form of cold gas from the observations of Storrie-Lombardi, McMahon & Irwin (1996) (the data converted to the ΛCDM cosmology was kindly provided by R. McMahon and C. Peroux). The observations shown should be considered a lower limit on the total mass of cold gas for several reasons. Some dusty DLAS may also be missed because their background quasars would be too dimmed to be included in an optically selected, magnitude-limited sample (Pei & Fall 1995). In addition, a significant amount of mass could be in the form of lower column density H I clouds, or in the form of molecular or ionized hydrogen.

We can see from Fig. 13 that the CE quiescent model (dashed line) overproduces the amount of cold gas by a factor that is perhaps a bit large to be explained by these effects, whereas the collisional starburst model (solid line) is more consistent with the data. The accelerated quiescent model (dotted line) is roughly 1σ lower than the data at $z \gtrsim 2$, and this may be a concern since all of the effects mentioned above will tend to increase this discrepancy. Note that the constraint provided by the DLAS is complementary to that provided by observations of LBGs. We could decrease the consumption of cold gas by decreasing the efficiency of star formation τ_*^0, but this would in turn produce fewer bright LBGs.

3.3.4 Metals

The metal content of the Universe is yet another tracer of the star formation history. The mean metallicity of the DLAS can be estimated from Zn II and Cr II absorption lines, which are not much depleted by dust (Pettini et al. 1997b). The metallicity of hot X-ray gas in clusters has been measured at $z \sim 0.3$ (Mushotzky & Loewenstein 1997) and $z = 0$ (Yamashita 1992; Butcher 1995). The metal content of the diffuse IGM is estimated from observations of the Lyman-α forest (references in figure caption). Figs 14 and 15 show the model predictions for the average metallicity of the entire Universe (total mass in metals divided by total mass of gas), and for the metallicity of stars, cold gas, hot gas in haloes, and diffuse photoionized gas that has been ejected from haloes. The cold gas lines from the models may be compared to the observations of DLAS, the hot gas lines to the observations of hot gas in clusters, and the diffuse gas lines to the metallicity of the higher column density Lyman-α forest. Note that the metals ejected from haloes in our models will probably not escape to large distances, but will remain in the vicinity of the haloes. Therefore our predictions for the metallicity of the 'diffuse' gas are probably more appropriately compared with the observational measurements in the higher column density Lyman-α forest.

The differences among the model results are not large. Overall, the collisional starburst model predicts the highest metallicities, the CE quiescent model the lowest. Aside from the DLAS metallicities, which are higher than the observations in all three models (see below), the largest discrepancy is in the Lyman-α forest, whose metallicity is too low in the CE quiescent model. The collisional starburst model predicts the flattest metallicity

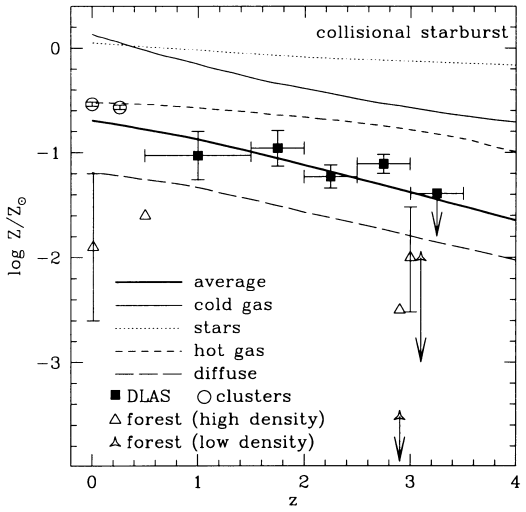
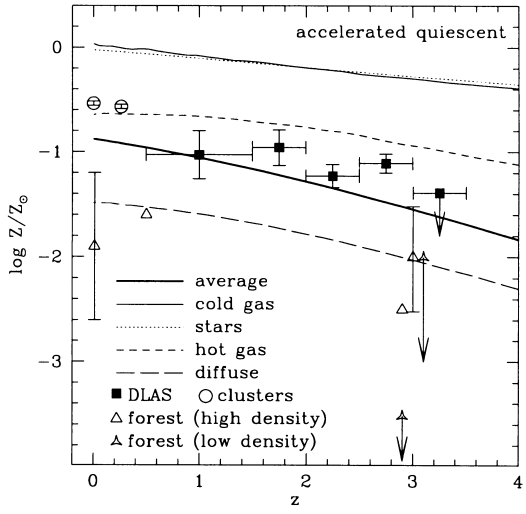


Figure 14. Metallicity as a function of redshift. The square symbols are from the measurements of Zn abundance in DLAS ([Zn/H_{DLA}]) from Pettini et al. (1997b). The dot at $z \sim 0.3$ is from measurements of Fe abundance in hot X-ray gas in clusters (Mushotszky & Loewenstein 1997), and the dot at $z = 0$ is the Fe abundance from local clusters (Yamashita 1992; Butcher 1995). Triangles show estimates of metallicity in the diffuse IGM from observations of the Lyman-α forest ($z \sim 0$: Shull et al. 1998; $z \sim 0.5$: Barlow & Tytler 1998; $z \sim 3$: Tytler & Fan 1994; Songaila & Cowie 1996; Rauch, Haehnelt & Steinmetz 1997; Lu et al. 1998). Lines show the collisional starburst model predictions of the mass-weighted mean metallicity in the form of stars, cold gas, hot gas, or diffuse (photoionized) gas, as labelled in the figure.



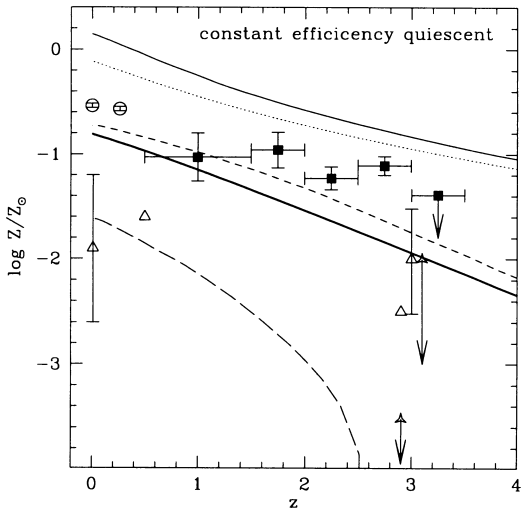


Figure 15. Same as Fig. 14, for the quiescent model.

evolution in all components, implying that the average metallicity of stars at $z \sim 3$–4 is already close to solar. The very small predicted evolution from $z = 0$ to $z \sim 0.3$ in the metallicity of the hot gas (residing mainly in large, cluster-sized haloes) in all three models is in good agreement with the observations of hot gas in clusters (Mushotszky & Loewenstein 1997).

As noted, the metallicity of cold gas in all models is systematically higher than that derived from DLAS. Since the metallicities of the cold gas and stars in galaxies are tightly coupled, it is difficult to imagine how one could make the metallicity of the cold gas much lower yet still match the observed solar metallicities of bright galaxies today. On the other hand, metallicities estimated from DLA surveys may systematically underestimate the true values for several reasons. First, dusty high-metallicity systems might dim any quasar in the line of sight enough to discourage observers from trying to obtain its spectrum (Pei & Fall 1995). Second, the metallicities of the outermost regions of local galaxies are often significantly smaller than their central metallicities. By a simple cross-section argument, there will be many more lines of sight passing through these metal-poor outskirts than the metal-rich inner regions of the galaxy. Finally, in the collisional starburst model, galaxies that have experienced a recent burst of star formation, accompanied by metal production, will have consumed nearly all of their cold gas and will not produce damped systems in absorption.

An interesting side effect of the collisional starburst scenario is that it predicts very early, efficient enrichment of the IGM and ICM by supernovae-driven winds. In our simple models, we have no difficulty in ejecting enough metals to pollute the IGM to the observed levels by $z \sim 3$. This is consistent with the strong metal-enriched outflows deduced from the spectra of observed LBGs (Pettini et al. 2000). This implies that a large fraction of the 'missing metals' at high redshift (Pagel 1998; Pettini 1999) may be in the form of hot gas in clusters or in the IGM. Similarly, enrichment occurs early in the accelerated quiescent model, again because galaxies with relatively small masses have high star formation rates. By contrast, in the CE quiescent model, most star formation takes place in very massive haloes, and gas and metals cannot escape from their deep potential wells. Significant pollution of the diffuse IGM therefore occurs only at much later times ($z \lesssim 1$).

3.4 Properties of galaxies at $z \sim 3$

Our results so far leave us with the conclusion that the CE quiescent model is in serious conflict with the observations. The best model seems to be the collisional starburst model, although the accelerated quiescent model is not strongly ruled out. In this section we concentrate on the collisional starburst model, and investigate the predicted properties of $z \sim 3$ galaxies in detail and compare them with the available observations. For this purpose, we have created 'mock-HDF' catalogues with the same volume as the HDF. Dust extinction is included using our usual recipe (see Section 2.4).

Fig. 16 shows the properties of galaxies in the redshift range $2 \lesssim z \lesssim 3.5$. The small dots show results as a function of the V_{606} (rest ~ 1600 Å) magnitude, for galaxies in a typical mock-HDF catalogue. Model galaxies have been selected with a flat selection function over the redshift range $2 \lesssim z \lesssim 3.5$. The histograms show the distribution of the same quantities for galaxies brighter than $V_{606} = 25.5$, to compare with ground-based samples. Because of the small volume of the HDF, we average over many random realizations to obtain the histograms.

Panel (a) of Fig. 16 shows the cold gas fraction $f_{\rm gas} \equiv m_{\rm cold}/(m_{\rm cold} + m_{\rm star})$. The distribution of $f_{\rm gas}$ is nearly uniform over the full range, with a median value of $\sim$0.5, and with many galaxies having values as high as $f_{\rm gas} = 0.8$–0.9. This is significantly higher than typical values for local disc galaxies,

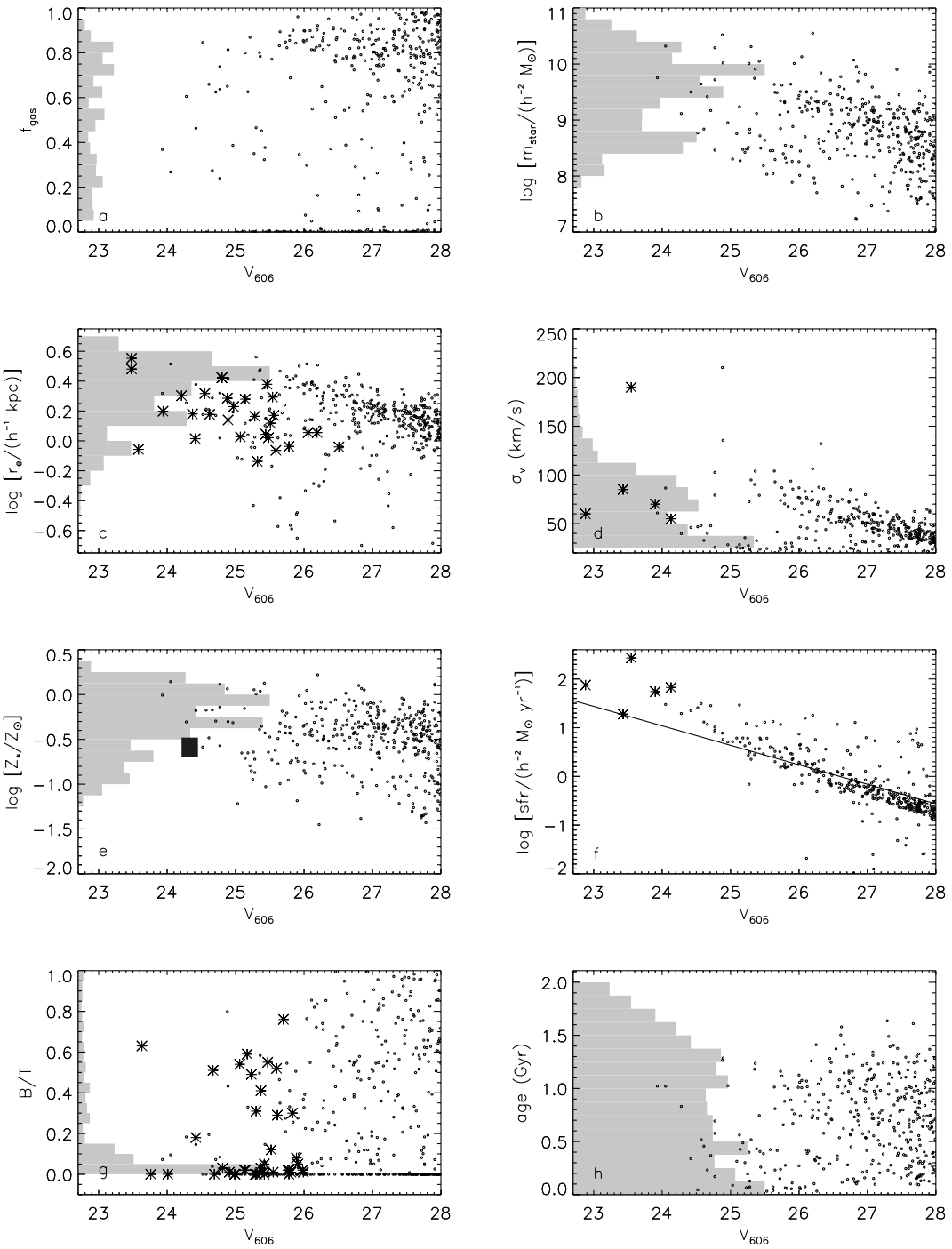


Figure 16. Small dots show various properties of galaxies in our mock-HDF catalogue, at $2 \leq z \leq 3.5$ as a function of the V_{606} (rest wavelength ~ 1500 Å) magnitude of the galaxy. The histograms show the distribution of galaxies in this redshift range that are brighter than $V_{606} = 25.5$. The quantities shown are: (a) gas fraction, (b) stellar mass, (c) half-light radius, (d) internal velocity dispersion, (e) stellar metallicity, (f) star formation rate, (g) bulge-to-total mass ratio and (h) mass-weighted mean stellar age. Large star-shaped symbols and the dark square in panel (e) show observational estimates of some of these quantities (see text).

which are around 0.1–0.25 (de Blok, McGaugh & van der Hulst 1996). We do not have direct observational measures of the gas content of the LBG population, but the high gas fractions we obtain are consistent with the factor of $\gtrsim$3 increase in the total cold gas content of the Universe ($\Omega_{\rm cold}$) from $z = 0$ to $z \sim 3$ from observations of damped Lyman-α systems, as demonstrated in Section 3.3.3.

Panel (b) shows the stellar masses of model LBG galaxies. Note that stellar masses range from 10^8 to $10^{11}\,h^{-2}\,{\rm M}_\odot$, on average one or two orders of magnitude smaller than the stellar masses of present-day $L*$ spirals and ellipticals. Thus, according to the collisional starburst model, observed LBGs are not the *already fully assembled* progenitors of present-day $L > L*$ galaxies (Giavalisco et al. 1996; Steidel et al. 1996a). It may well be the case that most of the stars in these LBGs would end up within bright galaxies at $z = 0$; however, there will be several generations of merging in the interim, and no simple one-to-one correspondence between the two populations. Accurate stellar masses for real LBGs are not yet available; however, the model stellar masses that we obtain are very compatible with the values calculated by Sawicki & Yee (1998) on the basis of multiband photometry of galaxies in the HDF.

Panel (c) compares the baryonic half-mass radii of model galaxies to actual data (see section 2.3 of SP for a discussion of how sizes are estimated in our models; see Mo, Mao & White 1999 for more detailed modelling of galaxy sizes). Typical values for the mock-catalogue galaxies are about 2 times smaller than those of nearby bright galaxies, further evidence that LBGs in the collisional starburst model are not fully assembled $L*(z = 0)$ galaxies. The radii of the model galaxies are in good agreement with the average half-light radii of LBGs observed in the HDF (starred points; Giavalisco et al. 1996; Lowenthal et al. 1997). Model radii at $z = 0$ are also in good agreement with the sizes of local galaxies, as we showed in SP. The significance of this agreement is less clear here, however, since the half-light radii, particularly in the UV, may well be considerably smaller than the baryonic half-mass radii that we model. However, radii measured in the observed near-IR (rest wavelength $\sim 4000\,$Å) from the NICMOS HDF are typically quite similar to those shown here (Dickinson et al. 1998).

Panel (d) shows observed and model linewidths. The velocity dispersions of observed LBGs can be estimated based on the widths of emission lines such as Hβ or O III. Emission lines have been detected for a few of the brightest LBGs, and the velocity dispersions derived from the observed linewidths are rather small: 50–90 km s^{-1} for four objects, and 180 km s^{-1} for one object (Pettini et al. 1998). The widths of the observed emission lines probably reflect the mass within about one effective radius. At such small radii, there is probably little contribution from dark matter. We therefore estimate the 1D velocity dispersion in our models using the expression

$$\sigma^2 = \frac{Gm}{cr_{\rm e}}, \tag{5}$$

where m is the total baryonic mass (cold gas and stars), $r_{\rm e}$ is the effective radius (we use the baryonic half-mass radius), and c is a geometry-dependent factor (Phillips et al. 1997), which we take to equal 6, corresponding to a hot component with a density $\rho \propto r^{-3}$ (Binney & Tremaine 1987). Although there is considerable uncertainty as to how to model the linewidths of LBGs, we obtain good agreement with the limited data available (Pettini et al. 1998; shown by the large star symbols).

With respect to Fig. 16(b), we remarked on the rather small masses of galaxies in the collisional starburst model. That is parallelled here by the modest velocity dispersions. Masses (and velocities) tend to be small in this model, because small objects are elevated to the level of visibility by starbursts. Such an effect occurs to a lesser extent in the accelerated quiescent model, but not at all in the CE quiescent model. This suggests that size indicators of all types (masses, radii, velocity dispersions) are a way to discriminate among models. In the accelerated quiescent model, for example, the shape of the distribution of dispersions is different – it is flatter and skewed towards larger σ_v but has a similar mean – while the CE quiescent model typically has the largest dispersions, ranging from 100 to 220 km s^{-1}, with a mean of about 140 km s^{-1}. More detailed modelling and additional data will provide important information on the baryonic masses of these objects.

Panel (e) presents metallicities for model galaxies. Little is known observationally about observed LBG metallicities, although wide variations in the strength of C IV absorption (see Steidel et al. 1996b, Lowenthal et al. 1997 and Trager et al. 1997) suggest that stellar metallicities have a broad range of values. The stellar metallicities of our model galaxies are consistent with this, ranging from about one-tenth solar to solar, with a mean of about one-half solar. This is also quite compatible with the estimate based on interstellar absorption lines in the lensed galaxy MS 1512-cB58 (Pettini et al. 2000; shown as the dark square in panel e). A weak metallicity–luminosity relationship is already in place at $z = 3$, with the brightest objects typically showing higher metallicities.

A tight correlation between rest-UV luminosity and instantaneous star formation rate is often assumed in order to estimate the former from the latter. The solid line in panel (f) shows the relation of this sort used by Madau, Pozzetti & Dickinson (1998). However, galaxies with complex star formation histories, particularly episodic ones, may show a non-negligible scatter from this relationship, as shown by our model galaxies. Dust extinction also increases the scatter in this relation and moves the brightest, most extinguished galaxies off it. The large star symbols in this panel show galaxies from the sample of Pettini et al. (1998), where the star formation rates have been estimated from Hβ emission lines. Star formation rates for the brightest galaxies in the collisional starburst model approach $100\,{\rm M}_\odot\,{\rm yr}^{-1}$, in agreement with observed values. The quiescent models predict far fewer galaxies with such high star formation rates, as reflected in Figs 4 and 5.

Panel (g) shows the mass ratio of stars in the bulge to the total stellar mass (B/T) for the model galaxies. We find that bright galaxies are biased towards low B/T ratios, indicating that the galaxies are disc-dominated. Observed B/T values (starred points) are also shown for bulge–disc decompositions of HDF galaxies, using the procedure described in Marleau & Simard (1998), where 'bulge' and 'disc' components were determined by fitting observed surface brightness profiles with a Sérsic or an exponential form. The availability of photometric redshifts allowed us to select galaxies in the redshift range of interest ($2.5 \leq z \leq 3.5$). These observations were compiled and provided to us by L. Simard and K. Wu. The resulting distribution of B/T with V_{606} magnitude looks very similar to the predictions of our models, including the weak apparent correlation of magnitude with B/T. However, our B/T values are weighted by mass, while observed values are weighted by light. This difference might bias the measured values in some unknown way. It is also interesting to note that about 20 per cent of the HDF galaxies [brighter than $m_{AB}(1500) \sim 26$]

have $B/T > 0.40$, similar to the fraction of early-type galaxies found in the nearby Universe.

The presence of an exponential profile does not mean that a classical disc is present, and similarly an observed $r^{1/4}$ profile does not imply the existence of an early-type galaxy in the classical sense. It is clear that traditional morphological definitions must be expanded and modified to usefully discuss the inhabitants of the high-redshift Universe. Many LBGs show a centrally concentrated $r^{1/4}$ core within a more diffuse envelope (Giavalisco et al. 1996), yet many also show pronounced substructure and signs of disturbance, even in the rest-visual NICMOS images (Lowenthal et al. 1997; Conselice et al. 1998). The collisional starburst scenario predicts that a measureable fraction of LBGs should show significant substructure and morphological peculiarity, and are quite unlikely to resemble classical disc galaxies. A quantitative statistical analysis of the morphology of observed LBGs could place important constraints on the collisional starburst scenario. We intend to pursue this topic in future work (preliminary results are presented in Somerville et al. 1999).

Accurate stellar ages of LBGs would provide a good test of models. However, the determination of ages from observed colours is complicated by the degenerate effect of reddening due to dust. Some observational papers discussing the LBG population have suggested that the total duration of star formation may be as large as 1 Gyr (Steidel et al. 1996b; Pettini et al. 1997a), based on the $\mathcal{R} - K$ colours of the objects. However, Sawicki & Yee (1998) conclude that the dominant stellar population of the LBGs is less than 0.2 Gyr old, with median ages of $\sim$10–36 Myr. This conclusion was based on their comparison of model SEDs to photometric data from five filter bands (V, I, J, H and K) spanning the Balmer break. The IR photometric data is helpful in breaking the age–extinction degeneracy, although some degeneracy remains. Panel (h) in Fig. 16 shows mass-weighted mean stellar ages for model galaxies at $z = 3$ as a function of apparent magnitude (note that the *luminosity*-weighted ages, which correspond more closely to the quantities estimated by Sawicki & Yee (1998), are considerably younger; see Section 3.4.1). The distribution of ages is fairly flat, with many objects having very young ages (indicating that they formed most of their stars in a recent burst) and some having ages close to the age of the Universe at the mean redshift of the sample.[1]

3.4.1 Dust extinction and colours

As described in Section 2.4, we have assumed a very simple recipe for dust extinction, in which the optical depth of a galaxy is a deterministic function of its intrinsic UV luminosity. Recall that the resulting extinction is also a function of the galaxy's inclination, which we chose randomly. Fig. 17 shows the effects of dust reddening on the model galaxies in our mock-HDF catalogue as a function of the unextinguished magnitude at rest wavelength 1500 Å. We convert the extinction in magnitudes (shown on the right-hand axis) to a rest frame colour excess $E(B - V)$ using the recipe given in Calzetti (1997). The solid histogram shows the distribution for the brightest (*attenuated* magnitude $m_{1500} < 25.5$) model galaxies. For comparison, we also show the distribution of rest $E(B - V)$ (provided to us by K. Adelberger) obtained from the spectral slopes of LBGs from the ground-based sample of Steidel et al. (1999), using the technique described by Meurer et al. (1999). A detailed discussion of dust extinction in observed LBGs is presented in Adelberger & Steidel (2000). Note that negative values of $E(B - V)$ are unphysical and result in part from the presence of Lyman-α emission in some of the spectra (not included in our modelling). In about half of the cases, Lyman-α is seen in absorption, and this broadens the $E(B - V)$ distribution redwards. Within these uncertainties, the distribution we obtain in our models is fairly similar to the observed one. This gives us some confidence that our simple dust model is at least roughly compatible with the best current observational estimates.

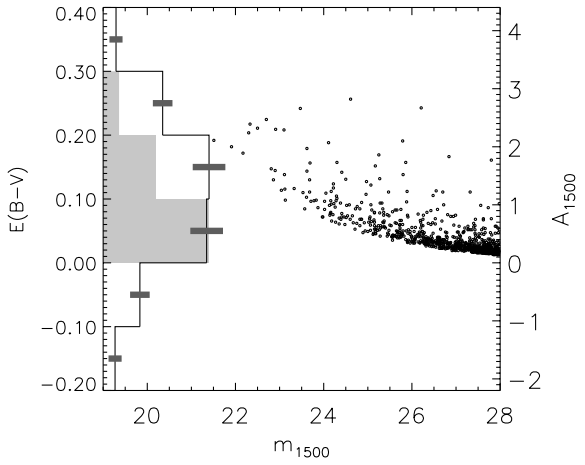


Figure 17. Extinction at rest wavelength 1500 Å for the model galaxies at $z = 3$, as a function of their *unattenuated* UV magnitude (see text for details of the dust recipe). The left-hand axis shows the extinction in terms of (rest-frame) $E(B - V)$, and the right-hand axis shows it in terms of magnitudes of extinction at 1500 Å. The scatter at fixed magnitude is due to the random inclination chosen for each galaxy. The shaded histogram shows the distribution of rest $E(B - V)$ for the bright (*attenuated* $m_{1500} <$ 25.5) galaxies in the models. The unshaded histogram shows the distribution of rest $E(B - V)$ for LBGs in the ground-based sample to similar magnitudes by (Adelberger & Steidel 2000).

We investigate the colours of our model galaxies, including the effects of dust, using the same recipe discussed above. Fig. 18 shows the $V_{606} - I_{814}$ and $I_{814} - K_s$ colours of galaxies in our mock-HDF (small dots), along with galaxies with spectroscopic (Lowenthal et al. 1997; Sawicki & Yee 1998) and photometric redshifts (Chen et al. 1999a; Lanzetta et al. 1999). It is immediately apparent that the model galaxies are too blue in $V_{606} - I_{814}$, by about 0.2 mag, while the $I_{814} - K_s$ colours are in reasonably good agreement with the data. One might be tempted to conclude from this that the starburst model produces galaxies with too young a stellar population; however, this is unlikely to be the explanation, as $I_{814} - K_s$ spans the Balmer break at this redshift and is a much better indicator of the age of the stellar population. The $V_{606} - I_{814}$ colour reflects the spectral slope in the far-UV ($\sim$1500–2000 Å), and, as we have discussed, this is strongly correlated with the far-IR excess, and hence with the amount of dust extinction (Meurer et al. 1999). This seems to indicate that we have underestimated the amount of dust extinction in our model galaxies, or that the true extinction curve is steeper than the Calzetti curve that we have assumed throughout – perhaps closer to an SMC curve. By comparing the reddening vectors shown in the figure, one can see that, had we used an SMC curve, both sets of colours would have been in quite good agreement with the data.

[1] The age of the Universe in this cosmology is 3.2 Gyr at $z = 2$, 2.1 Gyr at $z = 3$, and 1.8 Gyr at $z = 3.5$.

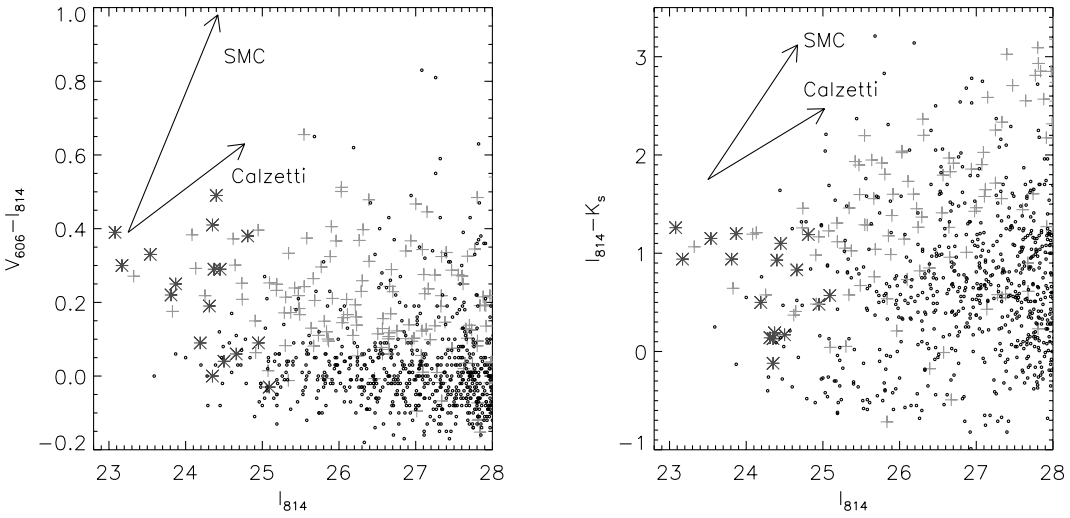


Figure 18. Observed $V - I$ (left) and $I - K$ (right) colours for galaxies at $z \sim 3$. The plus symbols show the colours of galaxies in the HDF with photometric redshifts in the range $2.0 \leq z \leq 3.5$ (Chen et al. 1999a; Lanzetta et al. 1999). Large star symbols show colours of HDF galaxies with spectroscopic redshifts (Lowenthal et al. 1997; Sawicki & Yee 1998). The small dots show the colours of model galaxies extracted from the mock-HDF catalogue, assuming a flat selection function in the range $2.0 \leq z \leq 3.5$. Dust extinction is included in the model galaxy colours using our usual recipe. The arrows show the reddening vector for an L_* galaxy at $z = 3$, assuming either a Calzetti or SMC extinction curve.

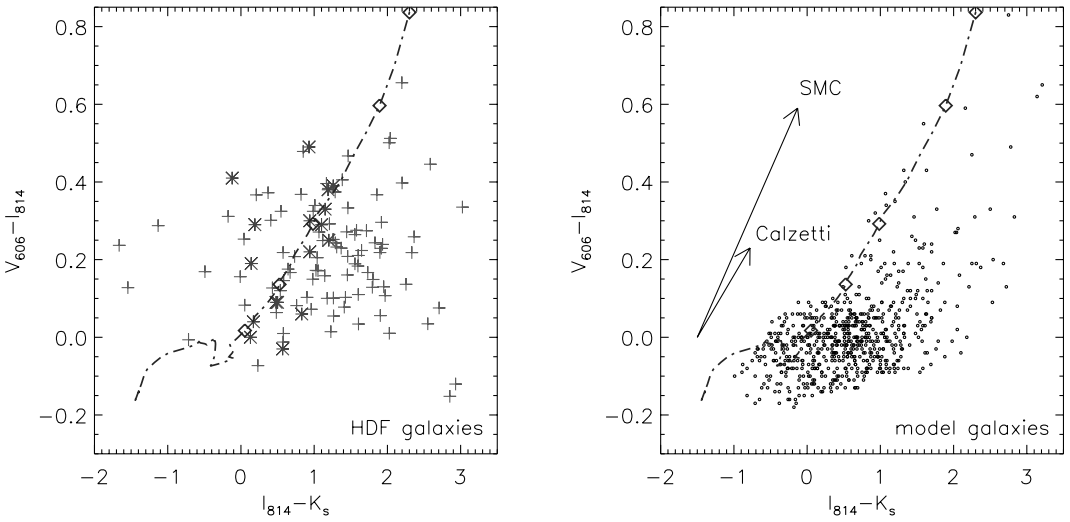


Figure 19. Two-colour diagram (observed $V - I$ versus $I - K$) for galaxies at $z \sim 3$. In the left-hand panel, large star symbols and plus symbols show the colours of HDF galaxies with spectroscopic and photometric redshifts (respectively) in the range $2.0 \leq z \leq 3.5$ (references in text). Small dots show the colours of model galaxies in the same redshift range from the mock-HDF catalogue (assuming a flat selection function from $2.0 \leq z \leq 3.5$), including dust extinction as usual. Dot-dashed lines show the unreddened colours of instantaneous burst populations of various ages from the solar metallicity Bruzual & Charlot models (GISSEL) at $z = 3$. The five diamond symbols mark the colours of single-burst ages of ~25, 50, 100, 250 and 320 Myr, from lower left to upper right. The arrows show the reddening vector for an L_* galaxy at $z = 3$, assuming a Calzetti- or SMC-type extinction curve. The length of the vector represents an extinction of a factor of 5 at 1500 Å.

A similar conclusion may be drawn from Fig. 19. The left-hand panel shows a two-colour diagram ($V_{606} - I_{814}$ versus $I_{814} - K_s$) for the observed HDF galaxies with spectroscopic (stars) and photometric (crosses) redshifts, as before. It is interesting to note the four or five photo-z galaxies with anomalously blue $I_{814} - K_s$ colours for their $V_{606} - I_{814}$ colours. A cross-check reveals that these galaxies are all fainter than $I_{814} = 25.5$, and therefore are not bright enough to be contained in the spectroscopic sample of Lowenthal et al. (1997). By comparing with the dot-dashed lines, which show the unreddened colours of a single-age stellar population from the GISSEL98 models (diamonds mark ages of ~25, 50, 100, 250 and 320 Myr from bottom left to top right), we see that their $I_{814} - K_s$ colours are characteristic of an extremely young (less than 25 Myr) stellar population, but the $V_{606} - I_{814}$ colours are considerably redder – perhaps these are very heavily extinguished starbursts? The right-hand panel shows where our model galaxies from the mock-HDF lie in this diagram. We do not see these anomalous objects in the models, even if an SMC extinction law is assumed, which could be an indication that our dust modelling is too simplistic (which would hardly be surprising), or that these objects are unusual, have been assigned incorrect redshifts, or have large photometric errors.

4 COMPARISON WITH PREVIOUS WORK

The results of this sort of modelling, particularly the redshift evolution, can be quite sensitive to the details of the recipes used to model the physical processes. This explains some of the differences between our results and those of other groups using similar techniques. For example, previous semi-analytic modelling of high-redshift galaxies (BCFL) gave a rather different picture of the star formation history of the Universe and the nature of the LBGs. BCFL concluded that observable LBGs form only in very massive haloes, $\gtrsim 10^{12} h^{-1} \mathrm{M}_{\odot}$, and implied that the dominant mode of star formation at all redshifts is quiescent. They state explicitly, and show in their fig. 7, that in their models 'most galaxies…never experience star formation rates in excess of a few solar masses per year.' This is clearly quite different from the results of our best-fitting models. These differences arise from the different ingredients that we have chosen to include in our models. BCFL made several assumptions that made the mass-to-light ratios of their haloes higher than ours. Based on our attempts to reproduce their results, we believe that some of the important differences are as follows.

(i) BCFL assumed primordial metallicity for all hot gas, reducing the efficiency of gas cooling (see fig. 2 of SP).

(ii) The explicit V_{c} dependence in their star formation recipe (Equation 1 of this paper; equations 2.5 and 2.10 of CAFNZ) makes star formation less efficient in small V_{c} haloes. This suppresses star formation at high redshift, when typical halo masses are smaller.

(iii) The strong V_{c} dependence of their supernovae feedback recipe (equations 2.8 and 2.11 of CAFNZ) further suppresses star formation in small haloes.

(iv) BCFL assume that only a fraction of the stars are luminous. In their fiducial model A, they assume that $f^{*}_{\mathrm{lum}} = 0.36$ ($\Upsilon = [f^{*}_{\mathrm{lum}}]^{-1} = 2.8$, in their notation), and in model G $f^{*}_{\mathrm{lum}} = 0.42$. This results in stellar mass-to-light ratios that are larger by a similar factor.

(v) Although merger-driven starbursts are included in the models of BCFL, they occur only in major mergers between satellites and central galaxies. This neglects the contribution from satellite–satellite mergers and minor mergers, which we have shown can be quite significant (see Fig. 11).

Fig. 20 shows the star formation history of the Universe as represented by the (almost) original Madau diagram (Madau et al. 1998), along with the results of BCFL model G (the closest of their models to our fiducial cosmology and IMF).[2] Also shown are our attempts to reproduce their results by adopting the assumptions listed above. Many differences in the details of the modelling remain, so it is not surprising that our results do not agree exactly, but they are sufficiently similar to give us confidence that we understand the main effects.

We also show the results of our CE quiescent model, the same model shown throughout the previous sections. It turns out that the results of model G of BCFL are very similar to those of our CE quiescent model. Although the efficiency of quiescent star formation is actually *lower* at early times in the models of BCFL [see points (ii) and (iii) above], starbursts in satellite–central major mergers were included [see point (v) above], and this contributes some additional bright galaxies. One can see from Fig. 20 and can further verify by comparing the luminosity functions shown in our Figs 5 and 7 with fig. 15 of BCFL, that, for our purposes, the models produce very similar results. Like our CE quiescent model, the models of BCFL produced just enough light and just enough bright galaxies at $z \sim 3$ *if dust extinction was neglected*. We now see that, when dust extinction is taken into account, these models produce a luminosity function that falls off too steeply at the bright end, and predict a steep fall-off in the number density of bright galaxies and the integrated luminosity density with redshift. This results in very few bright galaxies at redshifts of $\gtrsim 5$, in sharp contrast to the collisional starburst model or the accelerated quiescent model.

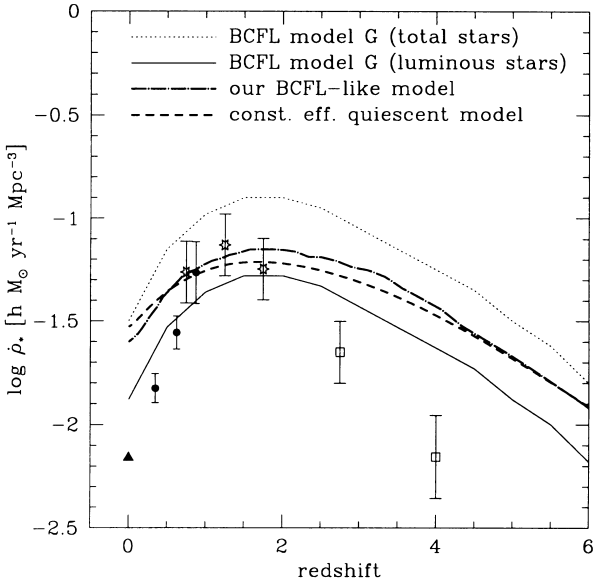


Figure 20. The (almost) original Madau diagram (Madau et al. 1998). The models shown are: model G of BCFL, our attempt to reproduce model G of BCFL, and our CE quiescent model. All three are seen to be similar. The dotted line is the curve shown in fig. 16 of BCFL. Our conversion of the data points from luminosity to star formation rate is different than the one used by BCFL, and is comparable to the star formation in *luminous* stars, shown for BCFL model G by the solid line.

As noted in Section 2, our accelerated quiescent star formation recipe is effectively very similar to the one usually used by the Munich group (e.g. KWG93). We have shown that this recipe produces fairly good agreement with many of the observations, but it does not produce enough very bright LBGs and also underpredicts the amount of cold gas at high redshift compared with observations of damped Lyman-α systems. In a recent paper, Kauffmann & Haehnelt (2000) reached a similar conclusion about the cold gas problem and adopted a similar solution. They write their quiescent star formation recipe as $\dot{m}* = \alpha m_{\mathrm{cold}}/t_{\mathrm{dyn}}$. As noted before, the dynamical time t_{dyn} scales approximately as $t_{\mathrm{dyn}} \propto (1+z)^{-3/2}$, so their adopted scaling of the parameter α with redshift as $\alpha(z) = \alpha(0)(1+z)^{-\gamma}$ with $\gamma = 1$–2 is very similar to our 'constant efficiency quiescent' recipe. They also include bursts in major satellite–central mergers, and find that these bursts are responsible for about two-thirds of the star formation at high redshift. They find that this model, which is similar to our favoured collisional starburst model, produces good results for the

[2] Note that in fig. 16 of BCFL, the data points shown were calculated using a different conversion from luminosity to star formation from ours (see Table A1 in our Appendix A, and table 5 of BCFL). In addition, BCFL plotted the 'total' star formation rate in the models, and corrected the observations assuming that only a fraction $[\Upsilon]^{-1}$ of the stars are luminous. We instead plot the star formation of *luminous* stars in all of the models.

redshift evolution of the space density of bright *quasars*. Taken together, this suggests a picture in which the brightest inhabitants of the high-redshift Universe, LBGs and quasars, may be produced by the same process: strong inflows induced by mergers.

It is also interesting to compare our results to those of cosmological *N*-body simulations including hydrodynamics and star formation. Katz, Hernquist & Weinberg (1999) performed a detailed study of the clustering of high-redshift galaxies in such simulations, and concluded that the observed clustering and number densities could be easily explained in a variety of CDM models. In this respect, their results are quite similar to the massive halo models discussed in Section 1. However, although their simulations included gas cooling and star formation, they assumed a monotonic relation between baryonic mass and UV luminosity. As we have already discussed, and as acknowledged by Katz et al., the effects of episodic star formation and dust could significantly complicate this relationship. A more direct look at the nature of the $z \sim 3$ galaxies in these simulations was taken by Davé et al. (1999); the star formation history of the galaxies was extracted and convolved with stellar population synthesis models to obtain estimates of their luminosities, and the effect of dust was considered. They then find that the luminosity function at $z \sim 3$ is in agreement with the observations when dust extinction is included. The galaxies brighter than 25.5 in their simulations have fairly large stellar masses ($\log m_{\rm star} \gtrsim 10.4$ for their value of $h_{100} = 0.65$) and contain a significant older stellar population; i.e., they are not predominantly young bursts. However, Davé et al. acknowledge that their simulations do not have sufficient resolution to properly model the collisional starburst mechanism that we have invoked here, so they cannot rule out the possibility that small starbursts could also be present. It is also worth noting that in these simulations, the star formation efficiency scales as the gas density to some power, which will give a similar behaviour to our 'accelerated quiescent' star formation recipe, and this may partially explain why they obtain sufficient numbers of LBGs without the need for an additional population of starbursts. The volume of their simulations is extremely small ($\sim 11\,h^{-1}$ Mpc on a side), far too small to probe the very bright end of the luminosity function, where we found the most pronounced differences between the collisional starburst model and the accelerated quiescent model.

5 DEPENDENCE ON MODELLING ASSUMPTIONS

This section discusses the sensitivity of our results to the cosmology, IMF, and stellar population modelling.

5.1 Cosmology and power spectrum

With improving observational constraints, the freedom to choose cosmological parameters as one wishes is rapidly dissappearing, and therefore we have chosen to focus on the popular ΛCDM cosmology. Any cosmology within the current observationally favoured regime of parameter space, namely $\Omega = 0.3$–0.5 (open or flat) and $\Gamma \simeq 0.2$, would give similar results. The old standard, cluster normalized SCDM, would also give similar results because of the large amount of power on small scales. Models with $\Omega = 1$ and realistic power spectra, such as τCDM, CHDM, or tCDM (see SP), show a stronger peak at $z \sim 1.5$ and a sharp decline in the star formation rate at higher redshift. In Fig. 21 we show the star formation history for a model in which we use the τCDM cosmology defined in SP and normalize to the local Tully–Fisher relation in our usual way. This figure should be compared to the corresponding Figs 8 and 9 for our standard ΛCDM cosmology. The τCDM model overproduces the luminosity density at low redshift ($z \lesssim 2$) and falls below the observations at high redshift ($z \sim 4$). If we normalize the model to fit the local UV-luminosity density, we find that there is not enough star formation at high redshift compared to the dust-corrected new Madau diagram. These models also violate the constraint on $\Omega_{\rm gas}$ from DLAS at high redshift. The results for models with similar power spectra such as CHDM or tCDM would be nearly identical.

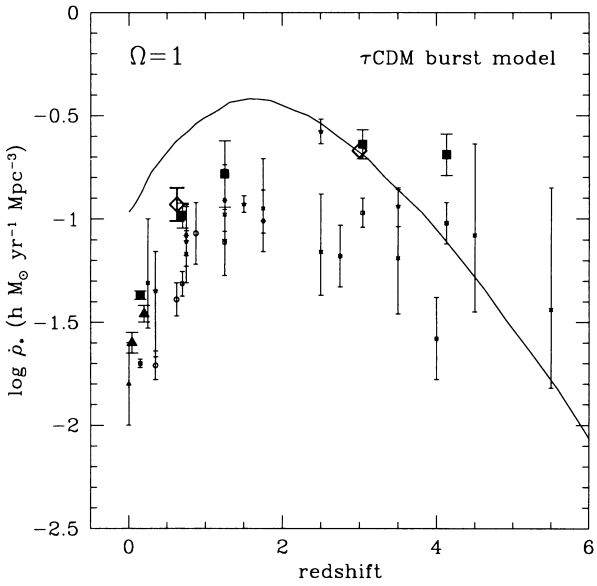


Figure 21. The star formation history in an $\Omega = 1$ universe. Small open and large filled points show the observations without and with (respectively) corrections for dust extinction (see Appendix A). The line shows a collisional starburst model in a τCDM cosmology, normalized to the Tully–Fisher relation at $z = 0$.

5.2 Stellar population synthesis

Most of the currently available observations of galaxies at $z \gtrsim 3$ are viewing light that is emitted in the far-UV ($\sim$1500 Å). This light is dominated by very young ($\lesssim$10 Myr), massive O- and B-type stars, which are relatively poorly understood and difficult to model. In addition, the stellar mass-to-light ratio depends on the metallicity and IMF.

In Fig. 22, we have made a comparison of the luminosity at 1500 Å (averaged over a top-hat with width 400 Å) of an instantaneous burst for a variety of the stellar population synthesis models that are publically available: GISSEL (Bruzual & Charlot 1993), PEGASE (Fioc & Rocca-Volmerange 1997), STARDUST (Devriendt et al. 1999), and STARBURST99 (Leitherer et al. 1999).[3] All of these models are for stars with solar metallicity and assume a Salpeter IMF, however, the STARBURST99 model is for a lower mass cut-off of 1 $M_\odot$, considerably larger than the value of 0.1 $M_\odot$ assumed by the other models. This is partially responsible

[3] The GISSEL models were provided to us by S. Charlot, the PEGASE models by M. Fioc, and the STARDUST models by J. Devriendt. The STARBURST99 models were downloaded from the public website http://www.stsci.edu/science/starburst99/.

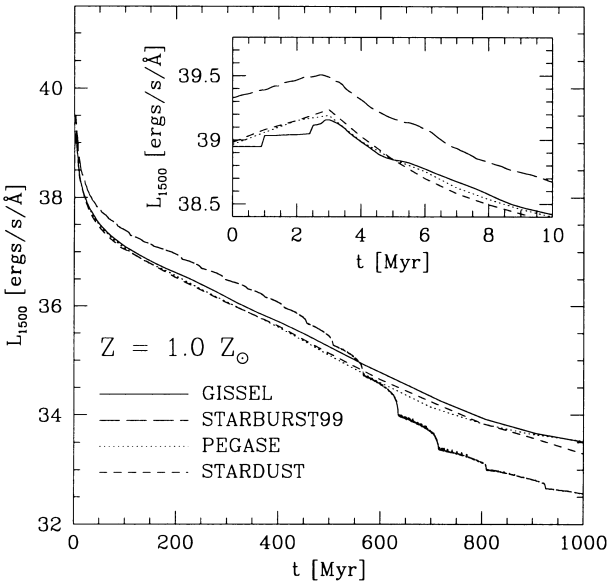


Figure 22. The luminosity of a single burst of mass $10^6\,M_\odot$ as a function of age for stellar population models from various groups (see text), all with solar metallicity and a Salpeter IMF.

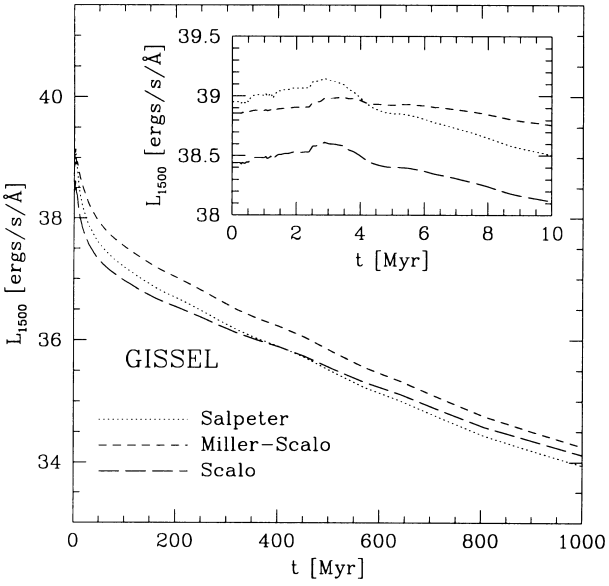


Figure 23. The luminosity of a single burst of mass $10^6\,M_\odot$ as a function of age for the GISSEL models with solar metallicity and various IMF. Retaining the Salpeter shape but changing the lower mass cut-off from 0.1 to $1\,M_\odot$ would result in shifting the luminosities up by about a factor of 3.

for the higher luminosity at young ages produced by STARBURST99 (see below). There is good agreement between the GISSEL, STARDUST, and PEGASE models. We find a similar level of agreement for populations with 0.2 and 0.4 solar metallicity.

Fig. 22 shows the dependence on the assumed IMF for the solar-metallicity GISSEL models. We see that the UV luminosity is quite sensitive to the IMF, in particular to the fraction of high-mass versus low-mass stars. The Scalo IMF gives results that are lower by about a factor of 3 at ages younger than 10 Myr. Similarly, changing the lower- or upper-mass cut-off would have an effect on the results. For example, if we drastically increased the lower mass cut-off to $1\,M_\odot$ instead of $0.1\,M_\odot$ and retained the Salpeter shape, the curve would shift up by about a factor of 3.[4] None of these changes are huge in the context of the theory, but evolutionary changes in the IMF as a function of redshift could induce a tilt in the Madau diagram, and thus affect comparison with observations. However, to salvage the CE quiescent model would require a total evolutionary change of a factor of 10 in mass-to-light ratio (see Fig. 9), which would stretch all the above effects to their limits.

The STARBURST99 models have been constructed to be particularly relevant to the very young stellar populations that we are interested in here, and use different stellar atmosphere models and a different approach to modelling mass-loss from the other three models mentioned above (see Leitherer et al. 1999). In Fig. 23 we show the dependence on metallicity in these models. For models with metallicities ranging from one-fifth solar to solar metallicity, the luminosity at 1500 Å is almost independent of metallicity for stars younger than 10 Myr. The luminosity becomes more sensitive to metallicity for older stars ($\gtrsim$200 Myr).

Note that we have neglected the contribution from nebular emission in our modelling. In the STARBURST99 models (Salpeter IMF with solar metallicity) nebular emission contributes about 15 per cent to the luminosity at 1500 Å at an age of 1 Myr.

[4] We base this on a model that we ran using the STARBURST99 website.

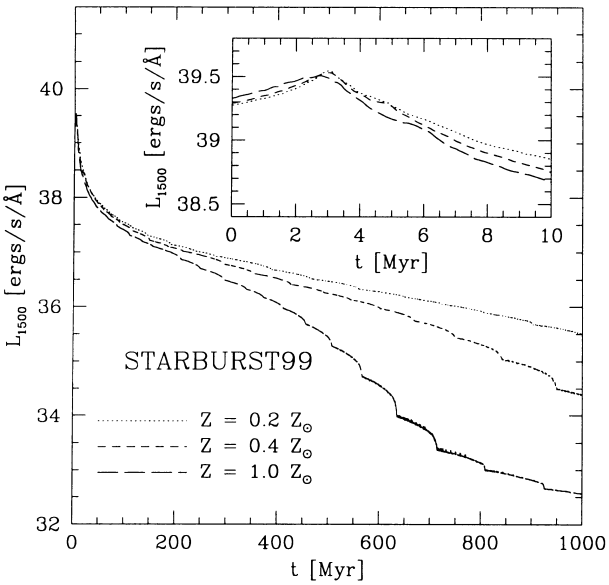


Figure 24. Same as Fig. 22, but for the STARBURST99 models with different metallicities. The far-UV luminosity does not depend much on the metallicity for the very young stars that dominate it.

5.3 Dust and reddening

Our treatment of dust and reddening is perhaps the least secure of our assumptions. This treatment falls into three broad categories: the mean absorption per typical galaxy, the variation in absorption versus wavelength (reddening curve), and the variation in absorption versus galaxy luminosity (brighter galaxies are more extinguished). None of these threaten the two main conclusions of this paper, namely that the collisional starburst model is consistent with present data on LBGs, and that star formation efficiency at high redshift must be higher than locally. With regard to the first

conclusion, we are claiming only consistency using the best information presently available. This has certainly been shown. With regard to the second point, even if we reduce the dust correction to zero, the CE quiescent model is inconsistent with the Steidel et al. (1999) data at $z \sim 4$ and preliminary detections of $z \gtrsim 5$ galaxies.

There are hints that some aspects of our dust treatment are in fact in error. For example, our model galaxies are systematically 0.2 mag too blue in $V - I$ (Fig. 18). Fixing this may necessitate changing the reddening curve to a steeper one closer in shape to the SMC curve; however, the overall amount of absorption would still remain roughly the same (see figure). Our third assumption (that brighter galaxies are more extinguished) is based in part on precisely the HDF galaxies we are modelling – after correction for extinction, is it precisely the reddest objects that are bolometrically brightest (Meurer et al. 1999).

Opinion seems to be converging on our first and most crucial assumption, the average absorption per typical galaxy. Given the observed $V - I$ colours of LBG galaxies, it is inconceivable that they have zero reddening. At the other extreme, dust absorptions 3 times larger than we have assumed would predict many more submillimetre sources than seen (Meurer et al. 1999). Tweaking the mean absorption within these boundaries would not greatly alter our picture.

On the other hand, a serious change to our assumed absorption treatment would be the discovery that mean absorption at a fixed luminosity declines strongly with redshift (the reverse variation is implausible). This would introduce a strong bias into all tests versus time, which is the most important probe of LBG models. Present data on the very distant Universe are simply too sparse to know for sure whether distant sources are systematically less obscured. Future submillimetre observations are needed to pin down this last, very important point.

6 SUMMARY AND CONCLUSIONS

The process of star formation is one of the largest uncertainties in galaxy formation modelling. We have explored three different recipes for star formation and the implications for observations of galaxies at high redshift. We investigated a model in which quiescent star formation has a constant efficiency, and galaxy–galaxy mergers trigger bursts of star formation. In this model, we find that most of the observable Lyman-break galaxies are collisional starbursts. We also considered two models in which star formation occurs only in a quiescent mode, with either constant or 'accelerated' efficiency.

We find that our collisional starburst model produces excellent agreement with the observed number density of bright galaxies as a function of redshift from $2 \leq z \leq 6$ when a reasonable amount of dust extinction is accounted for. The accelerated quiescent model produces nearly the same number density of bright galaxies as the starburst model over this redshift range. In contrast to the CE quiescent model, which predicts that very few bright galaxies would be in place at $z \gtrsim 4$, the burst model and the accelerated quiescent model make the tantalizing prediction that the comoving number density of bright galaxies at redshifts 5 and 6 is nearly as high as at $z \sim 3$. While this is consistent with the numbers of galaxies in the HDF-N and HDF-S with photometric redshifts in this range (Chen et al. 1999a; Lanzetta et al. 1999), and some spectroscopic confirmations of very high-redshift galaxies exist, secure estimates of the number densities of galaxies at these very high redshifts may have to wait for NGST and results from the new generation of infrared spectrographs on large ground-based telescopes.

A characteristic feature of the collisional starburst model is that the *unextinguished* rest-UV luminosity function at $z \sim 3$ is much flatter than a Schechter function at the bright end. Our prediction of the intrinsic luminosity function is in excellent agreement with the dust-corrected luminosity function of observed LBGs calculated by Adelberger & Steidel (2000). Conversely, if we introduce the effects of dust into the models using a simple, empirically motivated model of differential dust extinction (Wang & Heckman 1996), we obtain very good agreement with the actual observed luminosity function. Both the accelerated quiescent and the CE quiescent models show a steeper decline at the bright end. The CE quiescent model produces reasonable agreement with the observed bright end of the luminosity function at $z \sim 3$ *if dust extinction is neglected*, but with a realistic correction for dust extinction, these models predict a cut-off at the bright end that is much too steep. The accelerated quiescent model also has the wrong *shape* compared to the dust-corrected luminosity function of Adelberger & Steidel (2000). This model therefore underpredicts the number density of the very brightest LBGs ($\mathcal{R}_{AB} < 22$ at $z \sim 3$).

The global star formation history of the Universe is reflected in the redshift evolution of the total densities of star formation, cold gas, stars, and metals. We have compiled a new Madau diagram, using recent observations and an observationally based dust correction. We find a broad consistency between the dust-corrected optical estimates of the star formation rate density and those obtained from far-IR observations at $z \sim 0.7$ from *ISO* (Flores et al. 1999) and sub-mm observations at $z \sim 3$ from SCUBA (Hughes et al. 1998). The shape of the new Madau diagram is quite different from the original one, with a more gradual rise in the star formation rate from $z = 0$ to $z \sim 2$, and a plateau thereafter. The shape of the Madau diagram produced by our models is quite sensitive to our adopted recipe for star formation. The collisional starburst model is in excellent agreement with the new dust-corrected data, and the accelerated quiescent model is also in reasonable agreement. The CE quiescent model shows a much steeper decline with redshift, and has too little star formation at high redshift compared with the new data.

Similarly, the collisional starburst model and the accelerated quiescent model form their stars at a fairly high redshift, and thus stellar ages are in agreement with observational constraints from 'fossil evidence' contained in the ages of stars in nearby galaxies (Renzini 1998). The CE quiescent model forms its stars too late. In addition, we show that the redshift evolution of the density of cold gas is consistent with constraints from DLAS (Storrie-Lombardi et al. 1996) in the collisional starburst model. The CE quiescent model may have too much cold gas at high redshift, while the accelerated quiescent model may have too little.

We investigate the predicted metal enrichment history of the Universe, and find that an interesting side-effect of the burst and accelerated quiescent scenarios is efficient pollution of the IGM and ICM at high redshift due to supernovae-driven outflows. This may help to explain the large covering factor implied by the high probability with which C IV and Mg II absorption are seen in QSO spectra whenever a bright galaxy at the appropriate redshift is $\lesssim 30\,h^{-1}$ kpc of the QSO line of sight (Steidel 1995). It also implies that a substantial fraction of the metals at high redshift ($z \sim 3$) are in the form of hot gas in haloes or in the diffuse IGM, suggesting a solution to the 'missing metals' problem (Pagel 1998;

Pettini 1999). When normalized to produce solar-metallicity stars in bright galaxies at redshift zero, with no introduction of additional free parameters, the burst model simultaneously predicts the correct observed metallicity of diffuse gas (Lyman-α forest) at $z \sim 3$, as well as the observed metallicity of hot gas in clusters at $z = 0$ to 0.3. It is natural to associate the cold interstellar gas in our model galaxies with the observed population of DLAS. Here an interesting discrepancy is that the cold gas in our models has a systematically higher metallicity than the DLAS, by about a factor of $\sim$3 at all redshifts. Moreover, as the constraints on the metallicities of very low redshift ($z \lesssim 1.0$) DLAS improve (Pettini et al. 1999), if this gas is indeed the reservoir for star formation, it becomes difficult to understand how it can become enriched to solar values over such a short time interval. This favours the explanation that the observed population of DLAS systematically underestimates the metallicity of the total cold gas in galaxies.

The collisional starburst model produces good agreement with observable individual properties of $z \sim 3$ galaxies, such as the half-light radii and internal velocity dispersions. Our models also predict that high-redshift galaxies should have cold gas fractions that are much larger than present-day galaxies: values of $f_{\rm gas} \equiv m_{\rm gas}/(m* + m_{\rm gas})$ of 0.5 to 0.9 are typical. The model galaxies at $z \sim 3$ have perhaps rather surprisingly high metallicities, from 1/3 solar to solar for bright galaxies. The stellar masses of the model LBGs range from 10^8–$10^{11}\,h^{-2}\,{\rm M}_\odot$, up to three orders of magnitude smaller than present-day $L*$ galaxies. If our models reflect the real Universe, this indicates that the majority of the LBGs are not the fully formed, direct progenitors of today's $L*$ galaxies, but must experience considerable growth (via accretion or merging) by $z = 0$ if they are indeed the progenitors of massive present-day galaxies. Another possible destiny for some of the lower mass LBGs ($\lesssim 10^9\,h^{-2}\,{\rm M}_\odot$) is that their stars, gas, and globular clusters are stripped as they fall into the potential well of a nearby massive dark matter halo, forming a Population II stellar halo such as that of the Milky Way (Trager et al. 1997). This resembles the picture of galactic stellar halo formation outlined in the classic paper by Searle & Zinn (1978), which is supported by much recent evidence (see Carney et al. 1996, Majewski, Munn & Hawley 1996 and Sommer-Larsen et al. 1997, and references therein).

Finally, we compared our results with previous work and discussed the effects that different model assumptions have on our results. In the models of BCFL, star formation was made inefficient in small galaxies by the combined properties of the recipes for star formation and supernovae feedback. This led to the suppression of star formation at high redshift, and a prediction of a strongly decreasing star formation rate density with redshift. We argued that our CE quiescent model gives overall very similar results to the BCFL models, and therefore the same conclusions will apply: namely, these models are not consistent with the recent data at $z \sim 3$ and $z \sim 4$ when the observationally favoured correction for dust extinction is taken into account. The star formation recipe usually used by Kauffmann et al. is very similar to our accelerated quiescent model. It is interesting, however, that Kauffmann & Haehnelt (2000) recently exchanged this recipe in favour of one that is effectively rather similar to our model with CE quiescent star formation plus collisional starbursts. They disfavoured the 'accelerated' type model both because they found, as we did, that too much gas was consumed to be consistant with the observations of DLAS at $z \gtrsim 2$, but also because it did not reproduce the observed redshift evolution of the space density of bright quasars. Thus the same mechanism, merger-driven inflows, may account for both high-redshift galaxies and quasars.

We showed that in cosmologies with $\Omega = 1$ and realistic power spectra (e.g., τCDM, CHDM, tilted CDM), the decline in the star formation rate at high redshift is too steep to be consistant with the dust-corrected data even in the burst model. However, in any cosmology with parameters close to the values favoured by a broad range of observations ($\Omega_0 = 0.3$–0.5), we would have obtained similar results to those presented here. We showed a comparison of the UV stellar mass-to-light ratio in a variety of different stellar population models, and showed that there is reasonably good agreement between the models produced by different groups. We showed that the UV luminosites would have been about 3 times lower had we assumed a Scalo instead of Salpeter IMF, and about 3 times higher had we assumed a larger value for the lower mass cut-off ($1\,{\rm M}_\odot$ instead of $0.1\,{\rm M}_\odot$). We also showed that according to these models, the UV luminosity in galaxies with active star formation is nearly independent of the metallicity of the stellar population.

We summarize our main conclusions as follows.

(1) The details of the recipes used to model star formation can have very large effects on the results of galaxy formation models (either semi-analytic or numerical), especially when redshift evolution is considered.

(2) Models in which the efficiency of star formation is constant with redshift are strongly inconsistent with observations at high redshift ($z \gtrsim 3$) when the effects of dust extinction are taken into account, and are inconsistent with higher redshift observations ($z \gtrsim 4$) even when dust is neglected. Thus a rather robust conclusion of this study is that the efficiency of star formation (i.e., the star formation rate per unit mass of cold gas in a galaxy) must increase with redshift.

(3) This increased efficiency can be accomplished in at least two physically plausible ways: either due to collisional starbursts or to the scaling of the star formation rate with dynamical time or gas surface density. Both models produce good agreement with most of the observations considered here.

(4) We favour the collisional starburst mechanism, because it gives better agreement with the shape of the $z \sim 3$ luminosity function at the very bright end, and it is in better agreement with the constraints on the density of cold gas at high redshift from DLAS. However, due to the remaining uncertainties in the modelling, we do not consider the accelerated quiescent model to be strongly ruled out.

The two successful models are based on completely different physical ideas, but are difficult to distinguish conclusively from the present observations. It is crucial to eventually determine which process is actually the dominant mode of star formation at various redshifts. In one case, by studying high-redshift galaxies we can expect to learn something about the merger rate and the efficiency of merger-driven inflows. In the other case, we expect to learn more about the masses and internal properties of discs.

Both scenarios have observational support and theoretical motivation. It may, of course, be that the true situation involves a combination of both scenarios; however, simply adding starbursts to our usual accelerated quiescent recipe does not provide a solution. This is because the accelerated quiescent star formation very rapidly consumes the cold gas supply at high redshift. The contribution from the burst mode remains small because of the decreased gas fractions, and the constraints on the cold gas at high redshift from DLAS are badly violated. Real

progress in this sort of modelling will only be made by replacing some of our very simple recipes with more detailed and physically motivated prescriptions. This will require a better understanding of star formation in both normal and starburst galaxies.

An important observational aspect of the LBGs, which we have not addressed at all in this paper, is their clustering properties. Clustering in the collisional starburst scenario cannot be accurately modelled with analytic approaches. Therefore we study this using high-resolution *N*-body simulations in a companion paper (Kolatt et al. 1999). One might hope to discriminate between the collisional starburst and accelerated quiescent scenarios using the clustering properties of galaxies, particularly the number of close pairs. However, we show in a forthcoming paper (Wechsler et al. 2000) that this is not possible. We find that both scenarios are consistent with the observed clustering properties at $z \sim 3$.

There are, however, some direct observational tests which are feasible in the near future and which may begin to discriminate between the two scenarios. The collisional starburst scenario predicts that there should be a relatively large population of bright, heavily extinguished galaxies with star formation rates of hundreds of solar masses per year. There are already preliminary indications from the SCUBA results that this population has been detected. Future sub-mm experiments with higher resolution and sensitivity will put stronger constraints on the actual numbers and the associated total star formation rates associated with these objects. Another way of distinguishing the scenarios is from the morphologies of galaxies at high redshift. In the collisional starburst picture, we expect many of the LBGs to appear highly disturbed and to contain significant substructure. In the accelerated quiescent scenario, we expect the galaxies to appear smaller and denser, but otherwise similar to normal local spirals. Again, there are preliminary indications that a large fraction of the observed LBGs do show strong distortion and substructure, even in rest visual bands (Conselice et al. 1998; Dickinson et al. 1998), but a more quantitative analysis is needed. Finally, the observational sample of LBGs with measured emission linewidths is growing rapidly. With better statistics and more detailed modelling, these data will also help to discriminate between the two scenarios.

ACKNOWLEDGMENTS

We thank Kurt Adelberger, Daniela Calzetti, Romeel Davé, Mike Fall, Mauro Giavalisco, Ken Freeman, Andrey Kravtsov, Max Pettini, Jason Prochaska, Marcin Sawicki and Chuck Steidel for useful discussions, and Carlton Baugh, Shaun Cole, Carlos Frenk, Cedric Lacey, Guinevere Kauffmann and Simon White for their feedback on earlier versions of this work and clarifications of their models. We also thank Caryl Gronwall, James Lowenthal, Jesus Gallego, Rafael Guzman, Drew Phillips and Scott Trager for help with interpreting the observations. We are grateful to Stephane Charlot for providing us with updated versions of the GISSEL models. We thank Ken Lanzetta, Hsiao-Wen Chen and Amos Yahil for providing us with the Stonybrook photometric redshift catalogues and for useful discussions of their observations. We also thank Kurt Adelberger, David Hogg, James Lowenthal, Richard McMahon, Lucia Pozzetti, Marcin Sawicki, Luc Simard and Katherine Wu for providing observations in electronic form, and Neil Trentham for pointing out some errors in an earlier version of the manuscript. RSS acknowledges support from a University Fellowship from the Hebrew University, Jerusalem, and a PPARC rolling grant at the IoA. This work was also supported by a NASA ATP grant and NSF grant AST-9529098 at UCSC. We also acknowledge the Institute for Theoretical Physics at UC Santa Barbara, where this paper was completed.

REFERENCES

Adelberger K., Steidel C., 2000, ApJ, submitted, astro-ph/0001126
Adelberger K., Steidel C., Giavalisco M., Dickinson M., Pettini M., Kellogg M., 1998, ApJ, 505, 18
Barlow T., Tytler D., 1998, AJ, 115, 1725
Barnes J., Hernquist L., 1996, ApJ, 471, 115
Baugh C., Cole S., Frenk C., Lacey C., 1998, ApJ, 498, 504 (BCFL)
Beyer W., 1987, CRC Standard Mathematical Tables. CRC Press, Boca Raton, Florida
Binney J., Tremaine S., 1987, Galactic Dynamics. Princeton Univ. Press, Princeton, New Jersy
Blain A. W., Smail I., Ivison R. J., Kneib J.-P., 1999, MNRAS, 302, 632
Bruzual A., Charlot S., 1993, ApJ, 405, 538
Butcher J., 1995, PhD thesisUniv. Leicester,
Calzetti D., 1997, AJ, 113, 162
Carney B., Laird J., Latham D., Aguilar L., 1996, AJ, 112, 668
Charlot S., Worthey G., Bressan A., 1996, ApJ, 457, 625
Chen H.-S., Fernandez-Soto A., Lanzetta K., Pascarelle S., Puetter R., Yahata N., Yahil A., 1999a, astro-ph/9812339
Chen H.-S., Lanzetta K., Pascarelle S., 1999b, Nat, 398, 586
Cole S., Aragón-Salamanca A., Frenk C., Navarro J., Zepf S., 1994, MNRAS, 271, 781 (CAFNZ)
Connolly A., Szalay A., Dickinson M., SubbaRao M., Brunner R., 1997, ApJ, 486, L11
Conselice C.,et al., 1998, American Astronomical Society Meeting, 193
Cowie L., Songaila A., Hu E., Cohen J., 1996, AJ, 112, 839
Cowie L., Songaila A., Barger A., 1999, AJ, 118, 603
Davé R., Gardner J., Hernquist L., Katz N., Weinberg D., 1999, Clustering at High Redshift. Rencontres Internationales de l'IGRAP, astro-ph/9910220
de Blok W., McGaugh S., van der Hulst J., 1996, MNRAS, 283, 18
Devriendt J., Guiderdoni B., Sadat R., 1999, A&A, 350, 381
Dey A., Spinrad H., Stern D., Graham J., Chaffee F., 1998, ApJ, 498, L93
Dickinson M., 1998, in M. Livio S. F., Madau P., eds, The Hubble Deep Field, astro-ph/9802064
Dickinson M. et al., 1998, American Astronomical Society Meeting, Vol. No. 193, p. 75.11
Fioc M., Rocca-Volmerange B., 1997, A&A, 326, 950
Flores H. et al., 1999, ApJ, 517, 148
Fukugita M., Hogan C., Peebles P., 1998, ApJ, 503, 518
Gallego J., Zamorano J., Aragón-Salamanca A., Rego M., 1995, ApJ, 455, L1
Giavalisco M., Steidel C., Macchetto D., 1996, ApJ, 470, 189
Giavalisco M., Steidel C., Adelberger K., Dickinson M., Pettini M., Kellogg M., 1998, ApJ, 503, 18
Governato F., Baugh C., Cole S., Lacey C., Quinn T., Stadel J., 1998, Nat, 392, 359
Gronwall C., 1998, in Thuan T., eds, Dwarf Galaxies and Cosmology. Editions Frontières, astro-ph/9806240
Hu E., McMahon R., Cowie L., 1999, ApJ, 522, 9
Hughes D. H. et al., 1998, Nat, 394, 241
Jing Y., Suto Y., 1998, ApJ, 494, 5
Katz N., Hernquist L., Weinberg D. H., 1999, ApJ, 523, 463
Kauffmann G., Haehnelt M., 2000, MNRAS, 311, 576
Kauffmann G., White S., Guiderdoni B., 1993, MNRAS, 264, 201 (KWG93)
Kauffmann G., Colberg J., Diaferio A., White S. D. M., 1999, MNRAS, 303, 188
Kennicutt R., 1998, ApJ, 498, 181
Kolatt T. S., Bullock J. S., Sigad Y., Kravtsov A. V., Klypin A. A., Primack J. R., Dekel A., 2000, preprint, astro-ph/0010222

Kolatt T. S. et al., 1999, ApJ, 523, 109
Lanzetta K., Chen H.-S., Fernandez-Soto A., Pascarelle S., Puetter R., Yahata N., Yahil A., 1999, in Weymann R., Storrie-Lombardi L., Sawicki M., Brunner R., eds, Photometric Redshifts and High-redshift galaxies. astro-ph/9907281
Leitherer C. et al., 1999, ApJS, 123, 3
Lilly S., Le Fèvre O., Hammer F., Crampton D., 1996, ApJ, 460, L1
Lowenthal J. et al., 1997, ApJ, 481, 673
Lu L., Sargent W., Barlow T., Rauch M., 1998, preprint, astro-ph/9802189
Madau P., Ferguson H., Dickinson M., Giavalisco M., Steidel C., Fruchter A., 1996, MNRAS, 283, 1388
Madau P., Pozzetti L., Dickinson M., 1998, ApJ, 498, 106
Majewski S., Munn J., Hawley S., 1996, ApJ, 459, L73
Marleau F., Simard L., 1998, ApJ, 507, 585
Meurer G., Heckman T., Calzetti D., 1999, ApJ, 521, 64
Meurer G., Heckman T., Lehnert M., Leitherer C., Lowenthal J., 1997, AJ, 114, 54
Mihos J., Hernquist L., 1994, ApJ, 425, 13
Mihos J., Hernquist L., 1995, ApJ, 448, 41
Mihos J., Hernquist L., 1996, ApJ, 464, 641
Mo H. J., Fukugita M., 1996, ApJ, 467, L9
Mo H., Mao S., White S., 1999, MNRAS, 304, 175
Mushotszky R., Loewenstein M., 1997, ApJ, 481, L63
Oke J., Gunn J., 1983, ApJ, 266, 713
Pagel B., 1998, in Davis J., Impey C., Phillipps S., eds, The Low Surface Brightness Universe, ASP Conf. Ser., Vol. Astron. Soc. Pac., San Francisco
Pascarelle S. M., Lanzetta K. M., Fernandez-Soto A., 1998, ApJ, 508, L1
Pei Y., Fall S., 1995, ApJ, 454, 69
Pettini M., 1999, in Walsh J., Rosa M., eds, Chemical Evolution from Zero to High Redshift, ASP Conf. Ser., Vol. Springer, Berlin, astro-ph/9902173
Pettini M., King D., Smith L., Hunstead R., 1997a, ApJ, 478, 536
Pettini M., Smith L., King D. L., Hunstead R., 1997b, ApJ, 486, 665
Pettini M., Kellogg M., Steidel C., Dickinson M., Adelberger K., Giavalisco M., 1998, ApJ, 508, 539
Pettini M., Ellison S., Steidel C., Shapley A., Bowen D., 1999, ApJ, 520, 456
Pettini M., Steidel C., Adelberger K., Dickinson M., Giavalisco M., 2000, ApJ, 528, 96
Phillips A. et al., 1997, ApJ, 489, 543
Pozzetti L., Madau P., Zamorani G., Ferguson H., Bruzual A., 1998, MNRAS, 298, 1133
Primack J., 2000, in Courteau S., Strauss M., Willick J., eds, Cosmic Flows: Towards an Understanding of Large-Scale Structure, ASP Conf. Ser., Vol. Astron. Soc. Pac., San Francisco, astro-ph/9912089
Rauch M., Haehnelt M., Steinmetz M., 1997, ApJ, 481, 601
Renzini A., 1998, in D'Odorico S., Fontana A., Giallongo E., eds, The Young Universe: Galaxy Formation and Evolution at Intermediate and High Redshift, ASP Conf. Ser., Vol. Astron. Soc. Pac., San Francisco
Sawicki M., Yee H., 1998, AJ, 115, 1418
Sawicki M. J., Lin H., Yee H. K. C., 1997, AJ, 113, 1
Searle L., Zinn R., 1978, ApJ, 225, 357
Shull J., Penton S., Stocke J., Giroux M., van Gorkom J. H., Lee Y.-H., Carilli C., 1998, AJ, 116, 2094
Somerville R., 1997, PhD thesis, Univ. California, Santa Cruz, http://www.fiz.huji.ac.il/~rachels/thesis.html
Somerville R., Kolatt T., 1999, MNRAS, 305, 1
Somerville R., Primack J., 1999, MNRAS, 310, 1087 (SP)
Somerville R., Rosenfeld G., Kolatt T., Dekel A., Mihos J., Primack J., 1999, in Combes F., Mamon G., Charmandaris V., eds, Dynamics of Galaxies: from the Early Universe to the Present, ASP Conf. Ser., Vol. astro-ph/9910346
Somerville R., Lemson G., Kolatt T., Dekel A., 2000, MNRAS, 316, 479, astro-ph/9807277
Sommer-Larsen J. et al., 1997, ApJ, 481, 775
Songaila A., Cowie L., 1996, AJ, 112, 335
Spinrad H., Stern D., Bunker A., Dey A., Lanzetta K., Yahil A., Pascarelle S., Fernandez-Soto A., 1998, AJ, 116, 2617
Steidel C., 1995, in Meylan G., eds, QSO Absorption Lines, Proceedings of the ESO Workshop, ASP Conf. Ser., Vol. Springer-Verlag
Steidel C., Hamilton D., 1992, AJ, 104, 941
Steidel C., Hamilton D., 1993, AJ, 105, 2017
Steidel C., Giavalisco M., Dickinson M., Adelberger K., 1996a, AJ, 112, 352
Steidel C., Giavalisco M., Pettini M., Dickinson M., Adelberger K., 1996b, ApJ, 462, L17
Steidel C., Adelberger K., Dickinson M., Giavalisco M., Pettini M., Kellogg M., 1998, ApJ, 492, 428
Steidel C., Adelberger K., Giavalisco M., Dickinson M., Pettini M., 1999, ApJ, 519, 1
Stern D., Spinrad H., 2000, PASP, 111, 1475
Storrie-Lombardi L., McMahon R., Irwin M., 1996, MNRAS, 283, L79
Sutherland R., Dopita M., 1993, ApJS, 88, 253
Trager S., Faber S., Dressler A., Oemler A., 1997, ApJ, 485, 92
Trentham N., Blain A., Goldader J., 1999, MNRAS, 305, 61
Tresse L., Maddox S., 1998, ApJ, 495, 691
Treyer M., Ellis R., Milliard B., Donas J., Bridges T., 1998, MNRAS, 300, 303
Tytler D., Fan X.-M., 1994, ApJ, 424, L87
van Breugel W., De Breuck C., Stanford S., Stern D., Röttgering J., Miley G., 1999, ApJ, 518, L61
Van Dokkum P., Franx M., Kelson D., Illingworth G., 1998, ApJ, 504, L17
Walker I., Mihos J., Hernquist L., 1996, ApJ, 460, 121
Wang B., Heckman T., 1996, ApJ, 457, 645
Wechsler R., Gross M., Primack J., Blumenthal G., Dekel A., 1998, ApJ, 506, 19
Weymann R., Stern D., Bunker A., Spinrad H., Chaffee F., Thompson R., Storrie-Lombardi L., 1998, ApJ, 505, L95
Wechsler R. H., Somerville R. S., Bullock J. S., Kolatt T. S., Primack J. R., Blumenthal J. R., Dekel A., 2000, preprint, astro-ph/0011261
Williams R. et al., 1996, AJ, 112, 1335
Yamashita K., 1992, in Tanaka Y., Koyama K., eds, Frontiers of X-ray Astronomy, ASP Conf. Ser., Vol. Universal Academy Press, Tokyo
Zwaan M., Briggs F., Sprayberry D., Sorar E., 1997, ApJ, 490, 173

APPENDIX A: HOW TO DRAW A MADAU DIAGRAM

The star formation rate per unit comoving volume as a function of redshift was first compiled from observations extending from redshift zero to a redshift of about 4 by Madau et al. (1996). The now-famous 'Madau diagram' sketched out a picture of the history of star formation, from a very early epoch when galaxies were perhaps first forming until the present day. It has become popular to add more and more points to this diagram. However, unfortunately, as different authors have added their own points, the calculation of the derived quantity, the *total star formation rate density*, from what is actually observed (luminosities of galaxies selected in some way) has not always been consistent. We therefore think it is timely to revisit the steps in calculating this derived quantity from the observations, and to compile a set of points that have been calculated in as consistent a way as possible. As most of the results quoted in the literature assume an Einstein–de Sitter cosmology, we also provide the conversion to other cosmologies.

To draw your own Madau plot, follow these steps.

(i) **Correct for incompleteness**

Observational samples are generally flux-limited, and thus the intrinsic luminosity of the faintest objects in the sample changes with redshift. In order to understand the true redshift dependence of the *total* luminosity density, one must correct the observations for incompleteness. This is most easily done by fitting a functional form (i.e., a Schechter function) to the luminosity function

obtained from the observations themselves. If the usual parameters of the Schechter function, $\phi*$, $L*$ and α, are given, then the total luminosity density is given by $\phi* L* \Gamma(2+\alpha)$. Here Γ is the usual Gamma function (Beyer 1987). Note that for values of α steeper than -1, faint galaxies contribute a substantial fraction of the total luminosity density, and therefore the results are quite sensitive to the faint-end slope, which is often poorly constrained. This step leads to the first source of inconsistency: different authors have assumed different values of α, or have integrated down to different lower limiting luminosities. For example, the $z \sim 3$ and $z \sim 4$ values of the star formation rate density quoted in Steidel et al. (1999) were obtained by integrating down to a lower luminosity of $0.1\,L*$. The values quoted in other works, such as Gallego et al. (1995) and Lilly et al. (1996), were obtained by integrating down to zero luminosity. Combining these points on a single plot (as many authors have done in recent works) is therefore misleading.

Table A1. Conversion factors from luminosity to star formation (SFR $= C \times L$), where SFR is the star formation rate in $M_\odot\,yr^{-1}$, and L is the luminosity in $erg\,s^{-1}$ for Hα, and the luminosity density is $erg\,s^{-1}\,Hz^{-1}$ for the other tracers (Madau et al. 1998). A standard Salpeter IMF has been assumed.

Tracer	conversion factor C
Hα	6.31×10^{-40}
L_{1500}	1.25×10^{-28}
L_{2800}	1.27×10^{-28}

(ii) **Convert to the desired cosmology**

Most observational references quote luminosity densities assuming an Einstein–de Sitter ($\Omega = 1$) cosmology. To convert from one cosmology to another, one must take into account two effects. The luminosity derived from a given apparent magnitude will change, as will the comoving volume derived from a given angular size and redshift range. Luminosities scale as

$$\log\left(\frac{L_{\rm new}}{L_{\rm old}}\right) = 2\log\left[\frac{d_L^{\rm new}(z)}{d_L^{\rm old}(z)}\right], \tag{A1}$$

where $L_{\rm old/new}$ and $d_L^{\rm old/new}(z)$ are the luminosity and luminosity distance at a given redshift in the old and new cosmologies. If we now define $f_V \equiv V_{\rm new}/V_{\rm old}$ as the ratio of the comoving volume in the new cosmology to that in the old cosmology, then the luminosity density scales as

$$\log\left(\frac{\rho_L^{\rm new}}{\rho_L^{\rm old}}\right) = 2\log\left[\frac{d_L^{\rm new}(z)}{d_L^{\rm old}(z)}\right] - \log[f_V(z)]. \tag{A2}$$

(iii) **Convert luminosity to star formation**

The usual tracers of star formation are the luminosity of nebular emission lines like Hα or O II, or the far-UV (1500–2800 Å) continuum. This observable quantity must be converted into a star formation rate. This conversion generally relies on stellar

Table A2. Incompleteness corrected star formation rate densities at various redshifts, using various tracers of star formation, for $\Omega = 1$ ($q_0 = 0.5$; EDS), $\Omega_0 = 0.3$, $\Omega_\Lambda = 0.7$ (ΛCDM) and $\Omega_0 = 0.3$, $\Omega_\Lambda = 0.0$ (OCDM). Units are $h\,M_\odot\,yr^{-1}\,Mpc^{-3}$.

Reference	tracer	redshift	$\log[\rho_{SFR}^{\rm EDS}]$	$\log[\rho_{SFR}^{\Lambda{\rm CDM}}]$	$\log[\rho_{SFR}^{\rm OCDM}]$	log error
Gallego et al. 1995	Hα	0.0	−1.80	−1.80	−1.80	0.2
Gronwall 1998	Hα	0.0425	−1.60	−1.62	−1.61	0.05
Treyer et al. 1998	L_{2000}	0.15	−1.70	−1.76	−1.72	0.02
Tresse & Maddox 1998	Hα	0.2	−1.46	−1.53	−1.49	0.04
Lilly et al. 1996	L_{2800}	0.35	−1.71	−1.83	−1.75	0.07
		0.625	−1.39	−1.56	−1.46	0.08
		0.875	−1.07	−1.26	−1.15	0.15
Flores et al. 1999	$L_{\rm FIR}$	0.625	−0.93	−1.10	−0.98	0.08
Cowie et al. 1996	L_{2000}	0.7	−1.31	−1.49	−1.39	0.06
		1.25	−1.11	−1.33	−1.22	0.15
Connolly et al. 1997	L_{2800}	0.75	−1.08	−1.26	−1.16	0.15
		1.25	−0.91	−1.13	−1.02	0.15
		1.75	−1.01	−1.25	−1.14	0.15
Madau et al. 1996	L_{1500}	2.75	−1.4	−1.65	−1.56	0.15
		4.0	−1.90	−2.16	−2.08	0.2
Madau, Pozzetti & Dickinson 1998	L_{1500}	2.75	−1.18	−1.43	−1.34	0.15
		4.0	−1.58	−1.84	−2.08	0.2
Steidel et al. 1999	L_{1500}	3.04	−0.97	−1.22	−1.13	0.07
		4.13	−1.02	−1.27	−1.2	0.1
Hughes et al. 1998	sub-mm	3.0	−0.67	−0.92	−0.83	0.16
Sawicki et al. 1997	L_{2800}	0.35	−1.34	−1.46	−1.39	(+0.19)(−0.32)
		0.75	−1.11	−1.29	−1.19	(+0.05)(−0.07)
		1.5	−0.93	−1.16	−1.05	0.04
		2.5	−0.58	−0.83	−0.73	0.06
		3.5	−0.94	−1.19	−1.10	(+0.08)(−0.1)
Pascarelle et al. 1998	L_{1500}	0.25	−1.31	−1.40	−1.34	(+0.31)(−0.22)
		0.75	−1.17	−1.35	−1.25	(+0.23)(−0.14)
		1.25	−0.98	−1.20	−1.09	(+0.24)(−0.12)
		1.75	−0.95	−1.19	−1.08	(+0.24)(−0.12)
		2.5	−1.16	−1.41	−1.31	(+0.28)(−0.21)
		3.5	−1.19	−1.44	−1.36	(+0.34)(−0.27)
		4.5	−1.08	−1.34	−1.26	(+0.44)(−0.37)
		5.5	−1.44	−1.70	−1.63	(+0.59)(−0.38)

population models and an assumed star formation history and IMF (see Section 5.2 for a discussion of the attendant uncertainties and a quantitative comparison of different stellar population models). We give a compilation of conversion factors for various tracers of star formation in Table A1. These are taken from Madau et al. (1998), assuming a Salpeter IMF. The use of different conversion factors is a second source of inconsistency in published results in the literature.

(iv) **Correct for dust extinction**

If the tracer of star formation is an optical or UV luminosity, then the effects of dust extinction may be non-negligible. In the original Madau plot, no attempt was made to correct for extinction due to the lack of knowledge at that time about the effect that dust was likely to have on the observations, particularly at high redshift. As we discussed in the main text (see Section 2.4), some observational estimates of the extinction in the UV (∼1500–2000 Å) are now available. These seem to indicate that the amount of extinction at this wavelength in nearby starburst galaxies is similar to that in the $z \sim 3$ LBGs, if the correlation between spectral slope and dust extinction remains the same. However, the amount of extinction depends strongly on luminosity, with instrinsically brighter (more rapidly star-forming) galaxies being more heavily extinguished. Therefore it is probably not a good idea to apply a fixed correction factor to galaxies of all luminosities, as we would in effect be doing if we first corrected the luminosity density for incompleteness, as described in step (i), and then multiplied by a fixed factor to correct for dust extinction. If this correction factor was derived from observations of bright galaxies, but a substantial fraction of the total luminosity density is contributed by faint galaxies, this will lead to an overestimate of the total, dust-corrected luminosity density.

We give a compilation of star formation densities derived from various tracers at various redshifts in Table A2, for three different cosmologies (see table caption). These values have been corrected for incompleteness by integrating over the entire luminosity function, and converted from luminosity to star formation rate using the conversion factors from Table A1, but no correction for dust extinction has been made. Fig. 8 shows the resulting Madau plot for the ΛCDM cosmology, and Fig. 21 that for an $\Omega_0 = 1$ cosmology.

We now select a subset of these observations, those which we believe to give the most robust estimates of the star formation rate density over a broad range of redshifts. These samples represent a reasonably large volume (unlike the results from the HDF), are based on spectroscopic redshifts (unlike the results based on the more uncertain photometric redshifts), and were observed in roughly the same rest waveband (1500–2000 Å), and so do not require large photometric extrapolations or interpolations. We calculate the integrated UV luminosity density at low redshift from the results of Treyer et al. (1998), from Cowie, Songaila & Barger (1999) at intermediate redshift, and from Steidel et al. (1999) at high redshift. The sample with the best constrained faint-end slope is the lowest redshift sample of Treyer et al., which has a derived slope $\alpha = -1.6$, exactly the same as the faint end slope at $z = 3$ derived from the combined ground-based and HDF samples of LBGs (Steidel et al. 1999). We see no particular reason that the UV luminosity function should flatten at intermediate redshifts, so we assume $\alpha = -1.6$ for the Cowie et al. sample as well, and calculate all the incompleteness corrections accordingly (we use the values of $\phi*$ and $L*$ quoted by Cowie et al. 1999 for a fixed $\alpha = -1.5$, but we actually use $\alpha = -1.6$ in computing the luminosity density).

The results of Steidel et al. (1999) indicate that there is a factor of 4.7 extinction at ∼1500 Å in LBGs brighter than $\mathcal{R} = 25.5$, which corresponds to $0.4\,L*$ at the mean redshift of the sample ($z \sim 3$) and for our assumed cosmology. We will assume that these results hold for all of the UV-selected samples at all redshifts. To obtain a conservative ('minimal') dust correction, we apply the factor of 4.7 correction to all galaxies brighter than $0.4\,L*$ (using the appropriate value of $L*$ at each redshift) and no correction for galaxies fainter that this limit, for each of the three samples (Treyer et al. 1998, Cowie et al. 1999, and Steidel et al. 1999). Note that according to our previous calculation of dust extinction as a function of luminosity using the Wang & Heckman (1996) scaling, this will underestimate the dust correction but not by much (see Fig. 4). Because we have assumed that the slope of the UV luminosity function is the same for all of these samples, this is equivalent to adding 0.33 in the log to the values tabulated in Table A2. To obtain a 'maximal' dust correction, we apply the factor of 4.7 to *all* galaxies in the sample. As we discussed in item (iv) above, this is likely to be an overestimate. Note too that if we integrate the new Madau diagram with the *fiducial* dust correction, the total mass of stars produced by $z = 0$ is in good agreement with the estimates of Fukugita et al. (1998). However, if we integrate the Madau diagram with the *maximal* dust correction, the total mass in stars is 5 times too large (see Fig. 12). This suggests that something in between the minimal and maximal dust correction is probably reasonable (although one could get around this conclusion by varying the IMF from the assumed Salpeter shape).

We plot the results for our fiducial cosmology in Fig. 9. For comparison we also plot the results at low redshift from recent Hα surveys, corrected for incompleteness and dust extinction by the original authors (Gronwall 1998; Tresse & Maddox 1998). These surveys are quite deep, so incompleteness corrections are less important, and because Hα is emitted at ∼6500 Å, the effects of dust are also less severe. These results agree well with the Treyer et al. (1998) point. There is also a pleasing consistency between the dust-corrected, UV-based results at intermediate redshift and the far-IR results based on *ISO* observations at $z \sim 0.7$ (Flores et al. 1999), and finally the high redshift UV-based results and the sub-mm results from SCUBA observations at $z \sim 3$ (Hughes et al. 1998).[5]

A very different star formation history emerges from this 'new' Madau plot. Instead of the steep rise from $z = 0$ to $z \sim 1.5$, there is a more gradual rise, and instead of the peak at $z \sim 2$ and fall-off at higher redshift, there is a plateau.

This paper has been typeset from a TEX/LATEX file prepared by the author.

[5] It should be noted, however, that deriving the star formation rate from the IR and sub-mm observations is a more complex operation. See the discussions and more detailed modelling by Blain et al. (1999) and Trentham, Blain & Goldader (1999).